```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:      July 13, 2004, 12:34:44 ; Search time 88 Seconds
                                       (without alignments)
                            12108.099 Million cell updates/sec

Title:       US-09-270-437D-5
Perfect score: 560
Sequence:    1 agggacggctgccgcaccgcc.........atttccttcaggtttaaaa 1708

Scoring table: OLIGO
             Xgapop  60.0 , Xgapext  60.0
             Ygapop  60.0 , Ygapext  60.0
             Fgapop   6.0 , Fgapext   7.0
             Delop    6.0 , Delext    7.0

Searched:    1279676 seqs, 31191318243 residues

Word size:   1

Total number of hits satisfying chosen parameters:        2457832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
 -MODEL=frame+_n2p.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool_p/US09270437/runat_13072004_121958_10114/app_query.fasta_1.1863
 -DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
 -MAXLEN=200000000 -USER=US09270437 @CGN_1_130 @runat_13072004_121958_10114
 -NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :   Published_Applications_AA.*
      1 : /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
      2 : /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
      3 : /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
      4 : /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
      5 : /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
      6 : /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
      7 : /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
      8 : /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
      9 : /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
     10 : /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
     11 : /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
     12 : /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
     13 : /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
     14 : /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
     15 : /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
     16 : /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
     17 : /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
     18 : /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

                            SUMMARIES

 %
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result
 No.   Score  Query  Match Length DB  ID                              Description
------ -----  ----- ----- ------ --- --------                        -----------
```

Result No.	Score	Query	Match	Length	Pred. No.	ID	Seq
1	261	46.6	261	577	15	US-10-313-986-500	Sequence 500, App
2	216	38.6	216	261	9	US-09-764-864-1114	Sequence 1114, Ap
3	166	29.6	166	577	9	US-09-873-637-2	Sequence 2, Appli
4	116	20.7	116	250	9	US-09-764-864-1532	Sequence 1532, Ap
5	49	8.8	49	49	9	US-09-873-637-18	Sequence 18, Appl
6	48	8.6	48	48	9	US-09-873-637-20	Sequence 20, Appl
7	47	8.4	47	47	9	US-09-873-637-17	Sequence 17, Appl
8	45	8.0	45	47	9	US-09-873-637-19	Sequence 19, Appl
9	36	6.4	36	70	14	US-10-117-982-472	Sequence 472, App
10	36	6.4	36	70	15	US-10-313-986-472	Sequence 472, App
11	36	6.4	36	422	16	US-10-408-765A-2088	Sequence 2088, Ap
12	36	6.4	36	579	9	US-09-735-705-176	Sequence 176, App
13	36	6.4	36	579	9	US-09-735-705-348	Sequence 348, App
14	36	6.4	36	579	9	US-09-850-716A-176	Sequence 176, App
15	36	6.4	36	579	9	US-09-850-716A-348	Sequence 348, App
16	36	6.4	36	579	9	US-09-897-778-176	Sequence 176, App
17	36	6.4	36	579	9	US-09-897-778-348	Sequence 348, App
18	36	6.4	36	579	9	US-09-897-778-446	Sequence 446, App
19	36	6.4	36	579	9	US-09-897-778-449	Sequence 449, App
20	36	6.4	36	579	10	US-09-466-396A-176	Sequence 176, App
21	36	6.4	36	579	12	US-10-007-700-176	Sequence 176, App
22	36	6.4	36	579	12	US-10-007-700-348	Sequence 348, App
23	36	6.4	36	579	12	US-10-007-700-446	Sequence 446, App
24	36	6.4	36	579	12	US-10-007-700-449	Sequence 449, App
25	36	6.4	36	579	14	US-10-117-982-176	Sequence 176, App
26	36	6.4	36	579	14	US-10-117-982-348	Sequence 348, App
27	36	6.4	36	579	14	US-10-117-982-446	Sequence 446, App
28	36	6.4	36	579	14	US-10-117-982-449	Sequence 449, App
29	36	6.4	36	579	14	US-10-117-982-480	Sequence 480, App
30	36	6.4	36	579	14	US-10-117-982-484	Sequence 484, App
31	36	6.4	36	579	15	US-10-313-986-176	Sequence 176, App
32	36	6.4	36	579	15	US-10-313-986-348	Sequence 348, App
33	36	6.4	36	579	15	US-10-313-986-446	Sequence 446, App
34	36	6.4	36	579	15	US-10-313-986-449	Sequence 449, App
35	36	6.4	36	579	15	US-10-313-986-480	Sequence 480, App
36	36	6.4	36	579	15	US-10-313-986-484	Sequence 484, App
37	36	6.4	36	586	9	US-09-850-716A-427	Sequence 427, App
38	36	6.4	36	586	9	US-09-897-778-427	Sequence 427, App
39	36	6.4	36	586	12	US-10-007-700-427	Sequence 427, App
40	36	6.4	36	586	14	US-10-117-982-427	Sequence 427, App
41	36	6.4	36	586	15	US-10-313-986-427	Sequence 427, App
42	36	6.4	36	589	15	US-10-313-986-486	Sequence 486, App
43	28	5.0	28	47	9	US-09-873-637-21	Sequence 21, Appl
44	24	4.3	24	70	14	US-10-117-982-473	Sequence 473, App
45	24	4.3	24	70	15	US-10-313-986-473	Sequence 473, App

```
                            ALIGNMENTS

RESULT 1
US-10-313-986-500
; Sequence 500, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
;  APPLICANT: Foy, Teresa M.
;  APPLICANT: McNabb, Andria
;  APPLICANT: Watanabe, Yoshihiro
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Wang, Tongtong
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C19
;  CURRENT APPLICATION NUMBER: US/10/313,986
;  CURRENT FILING DATE: 2002-12-04
;  NUMBER OF SEQ ID NOS: 560
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 577
;  LENGTH: 577
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-10-313-986-500

Alignment Scores:
```

```
Pred. No.:              8.27e-241      Length:              577
Score:                  261.00         Matches:             441
Percent Similarity:     99.55%         Conservative:        0
Best Local Similarity:  99.55%         Mismatches:          0
Query Match:            46.61%         Indels:              2
DB:                                    Gaps:                1

US-09-270-437D-5 (1-1708) x US-10-313-986-500 (1-577)

Qy    38  GCCATCATGAAGCTGAAGGGCCACCAGTTGAGAACCATGCCCTGAAGGTCTCCTACATC       97
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   135  AlaIleMetLysLeuLysGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIle      154

Qy    98  CCCGATGAGCAGATAGCACAGGAGCCTGAGAATGGGCGCCCAGGGGCCTTTGGCTTCCGG      157
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   155  ProAspGluGlnIleAlaGlnGlyProGluAsnGlyArgArgGlyAlaGlyPheGlySerArg   174

Qy   158  GGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGCGGGGGCCCCAGCCAAGCAGCAGAGTG      217
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   175  GlyGlnProArgGlnGlySerProValAlaAlaGlyAlaProAlaAlaLysGlnGlnArgVal   194

Qy   218  GACATCCCCTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAG      277
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   195  AspIleProLeuArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGlu      214

Qy   278  GGGGCCACCATCCGCAACATCACAAAACAGACAGACCAGTCCAAGATAGAGTGTCATAGGAAG     337
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   215  GlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleGluAspValHisArgLys   234

Qy   338  GAGAACGCAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTCC      397
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   235  GluAsnAlaGlyGlnLeuLysLysAlaIleSerValHisSerThrProGluGlyCysSer      254

Qy   398  TCCGGCTTGTGAAGATGATCTTGGAGATTATGCATAAGAGGCTAAGGACACCAAAAACCGGCT   457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   255  SerAlaCysLysMetIleLeuGluIleMetHisLysGluLeuArgThrLysThrArgLeu      274

Qy   458  GACGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTGTAGGGGTGTCTCATTGGCAAG      517
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   275  AspAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLys   294

Qy   518  GAAGGACGGAACCTGAGAAGTAGAGCAAGATACCGAGACAAAATACCATCTCCTCG      577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   295  GluGlyArgAsnLeuLysLysValGluGlnAspThrGlnThrLysIleThrIleSerSer      314

Qy   578  TTGCAAGACCTCACCCTTACACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAG      637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   315  LeuGlnAspLeuThrLeuTyrAsnProGluThrIleThrValLysGlyAlaIleGlu      334

Qy   638  AATTGTTGCAGGGCCGAGCAGGAAATAATGAAGAAAGTTCGGGAGGCCCATAGGAGAATGAT    697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   335  AsnCysCysArgAlaGluGlnGluIleMetLysLysValArgGluAlaHisArgArgMetMet   354

Qy   698  GTGGGCTGCCATGAGC------TCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAGGT    751
          ||||||||||||||      |||||||||||||||||||||||||||||||||||||||||
Db   355  ValAlaAlaMetSerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAlaValGly      374

Qy   752  CTTTTCCCAGCTTCATCCAGCCAGTCCCGCGCCCTCCCAGCAGCGTTACTGGGGCTGCT      811
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   375  LeuPheProAlaSerSerSerAlaValProProProSerValThrGlyAlaAla      394

Qy   812  CCCTATAGTCCTTTATGCAGGTCCCCAGCAGCAGGAGATGGTGCAGGTGTTTATCCCCGCC      871
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   395  ProTyrSerSerPheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAla      414

Qy   872  CAGGGAGCTGGGGCCATCATCGGCAAGAGGCAGCACATCAAACAGCTTCTCCGGTTT      931
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   415  GlnAlaValAlaGlyAlaIleIleGlyLysGlnHisIleLysGlnLeuSerArgPhe      434

Qy   932  GCCAGCCGCCTCCATCCAAGATTGCACACTCAAACACTGACTCCAAAGTTCGTATGGTT      991
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   435  AlaSerArgAlaSerIleLeuIleLysGlnHisSerLysLeuProThrAspGlnAlaThrArgMet   454
```

```
Qy   992  ATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGGAGAATCTATGCAAACTC     1051
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   455  IleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeu      474

Qy  1052  AAGGAGGAGAACTTCTTGGTCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTGTG     1111
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   475  LysGluGluAsnPheLeuValGlnGlyGlyValLysLeuGluThrHisIleArgVal      494

Qy  1112  CCAGGATCAGCAGCTGGCCCGGTCATTGGAAAACGGTGGAAAACGGTGAACGAGTTGCAG     1171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   495  ProAlaSerAlaAlaGlyProValIleGlyLysGlyGlyGlnThrValAsnGluLeuGln      514

Qy  1172  AATTTGACGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGAACGACCAG     1231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   515  AsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsnAspGln      534

Qy  1232  GTCATCGTGAAAATCATCGGACATTCTTATGCCGATGGCTCACGGAGGATCCGA     1291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   535  ValIleValLysIleIleGlyHisSerTyrAlaSerGlnMetAlaGlnArgArgIleArg      554

Qy  1292  GACATCCTGCCCCAGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACCAGGCCCAGGCA     1351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   555  AspIleLeuAlaGlnLeuSerSerSerIleArgArgAspArgValThrArgProArgGln      574

Qy  1352  CGGAGGAAG     1360
          |||||||||
Db   575  ArgArgLys     577
```

RESULT 2
; US-09-764-864-1114
; Sequence 1114, Application US/09764864
; Patent No. US2002013275A1
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PT223
;   CURRENT APPLICATION NUMBER: US/09/764,864
;   CURRENT FILING DATE: 2001-01-17
;   Prior application data removed - consult PALM or file wrapper
;   NUMBER OF SEQ ID NOS: 1792
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 1114
;   LENGTH: 261
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-764-864-1114

```
Alignment Scores:
Pred. No.:              1.26e-197      Length:              261
Score:                  216.00         Matches:             216
Percent Similarity:     100.00%        Conservative:        0
Best Local Similarity:  100.00%        Mismatches:          0
Query Match:            38.57%         Indels:              0
DB:                                    Gaps:                0

US-09-270-437D-5 (1-1708) x US-09-764-864-1114 (1-261)

Qy   713  TCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGC      772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    46  SerHisLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAlaSerSerSer      65

Qy   773  GCAGTCCCGCGCCCTCCCAGCAGCGTTACTGGGGCTGCTCCCTATAGTCCTTTATGCAG      832
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    66  AlaValProProProSerValThrGlyAlaAlaProTyrSerSerPheMetGln      85

Qy   833  GCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGCCCAGGGAGCTGGGGCCATCATC      892
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    86  AlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValAlaGlyAlaIleIle   105

Qy   893  GGCAAGAGGGGCAGCACATCAAACACTGACTCCAAAGTTCGTATGGTT      952
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   106  GlyLysLysGlyGlnHisIleLysGlnHisSerLysLeuSerArgLysAlaSerIleLeuIle   125
```

RESULT 3
US-09-873-637-2
; Sequence 2, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-2

Alignment Scores:
Pred. No.:   577
Score: 9.89e-150 166.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 29.64%
DB: 9

RESULT 4
US-09-764-864-1532
; Sequence 1532, Application US/09764864
; Patent No. US2002013275A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1532
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (134)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1532

Alignment Scores:
Pred. No.:   250
Score: 9.44e-102 116.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 20.71%
DB: 9

```
Db    215 HisPheTyrAlaSerGlnMetAlaGlnHisMetAlaGlnMetAlaGlnGlnAsnProAlaGlnIleLeuAlaGlnGlnValLys 234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1313 CAGCAGCATCAGAAGGGACAGAGTAACCAGGCACCAGGCCACGGAGGAAG 1360
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    235 GlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaAlaArgArgLys 250

RESULT 5
US-09-873-637-18
; Sequence 18, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-18

Alignment Scores:
Pred. No.:                   1.97e-37        Length:          49
Score:                       49.00           Matches:         49
Percent Similarity:          100.00%         Conservative:    0
Best Local Similarity:       100.00%         Mismatches:      0
Query Match:                 8.75%           Indels:          0
                                             Gaps:            0

US-09-270-437D-5 (1-1708) x US-09-873-637-18 (1-49)

Qy    476 ATCCTGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAGGAGACGGAACCTGAAG 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 IleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyIleGlyArgAsnLeuLys 20

Qy    536 AAGGTAGAGCAAGATACCGAGACAAAATCACCATCTCCGTTGCAAGACCTTACCCTT 595
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 LysValGluGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeuThrLeu 40

Qy    596 TACAACCCTGAGAGGACCACCATCACTGTG 622
          ||||||||||||||||||||||||||||||
Db     41 TyrAsnProGluArgThrIleThrVal 49

RESULT 6
US-09-873-637-20
; Sequence 20, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-20

Alignment Scores:
Pred. No.:                   1.8e-36         Length:          48
Score:                       48.00           Matches:         48
Percent Similarity:          100.00%         Conservative:    0
Best Local Similarity:       100.00%         Mismatches:      0
Query Match:                 8.57%           Indels:          0


DB:                          9               Gaps:            0

US-09-270-437D-5 (1-1708) x US-09-873-637-20 (1-48)

Qy   1103 ATACGTGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGTGGAAAAACGGTGAAC 1162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 IleArgValProAlaSerAlaAlaGlyArgValIleGlyLyGlyLysThrValAsn 20

Qy   1163 GAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAG 1222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 GluLeuGlnAsnLeuThrAlaAlaAlaGluValValValProArgAspGlnThrProAspGlu 40

Qy   1223 AACGACCAGTCATCGTGAAAATC 1246
          |||||||||||||||||||||||
Db     41 AsnAspGlnValIleValLysIle 48

RESULT 7
US-09-873-637-17
; Sequence 17, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-17

Alignment Scores:
Pred. No.:                   1.64e-35        Length:          47
Score:                       47.00           Matches:         47
Percent Similarity:          100.00%         Conservative:    0
Best Local Similarity:       100.00%         Mismatches:      0
Query Match:                 8.39%           Indels:          0
                                             Gaps:            0

US-09-270-437D-5 (1-1708) x US-09-873-637-17 (1-47)

Qy    233 CTCCTGGTGCCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGAGGGGGCCACCATCCGC 292
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 LeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyIleAlaThrIleArg 20

Qy    293 AACATCACAAAACAGACCCAGTCCAAGATAGAGTGCATAGAAGGAGAAGAACCAGAGTGCA 352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AsnIleThrLysGlnThrGlnSerLysIleGluLysIleAspValHisArgLysAsnAlaGlyAla 40

Qy    353 GCTGAAAAAGCCATCAGTGTG 373
          |||||||||||||||||||||
Db     41 AlaGluLysAlaIleSerVal 47

RESULT 8
US-09-873-637-19
; Sequence 19, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 47
; TYPE: PRT
```

; ORGANISM: Mus musculus
US-09-873-637-19

Alignment Scores:
Pred. No.:
Score: 1.35e-33
Percent Similarity: 45.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 8.04%

Length: 47
Matches: 45
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 9

US-09-270-437D-5 (1-1708) x US-09-873-637-19 (1-47)

Qy   857 GTGTTATCCCGCCCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGGCAGCACATCAAA 916
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ValPheIleProAlaGlnAlaValGlyAlaIleIleGlyLysLysGlyGlnHisIleLys 20

Qy   917 CAGCTCTCCCGGTTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCC 976
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 GlnLeuSerArgPheAlaSerAlaSerIleLysIleAlaProProGluThrProAspSer 40

Qy   977 AAAGTTCGTATGGTT 991
          |||||||||||||
Db   41 LysValArgMetVal 45

RESULT 9
US-10-117-982-472
; Sequence 472, Application US/10117982
; Publication No.US2003013438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 472
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-472

Alignment Scores:
Pred. No.:
Score: 5.48e-25
Percent Similarity: 36.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6.43%

Length: 70
Matches: 36
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 14

US-09-270-437D-5 (1-1708) x US-10-117-982-472 (1-70)

Qy   254 GTGGGTGCCATTATTGGCAAGGAGGGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 37

Qy   314 TCCAAGATAGACGTGCATAGGAAGGAAGACGCAGGTGCAGCTGAAAAA 361
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   38 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 53

RESULT 10

US-10-313-986-472
; Sequence 472, Application US/10313986
; Publication No.US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 472
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-472

Alignment Scores:
Pred. No.:
Score: 5.48e-25
Percent Similarity: 36.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6.43%

Length: 70
Matches: 36
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 15

US-09-270-437D-5 (1-1708) x US-10-313-986-472 (1-70)

Qy   254 GTGGGTGCCATTATTGGCAAGGAGGGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 37

Qy   314 TCCAAGATAGACGTGCATAGGAAGGAAGACGCAGGTGCAGCTGAAAAA 361
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   38 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 53

RESULT 11
US-10-408-765A-2088
; Sequence 2088, Application US/10408765A
; Publication No.US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2088
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2088

Alignment Scores:
Pred. No.:
Score: 4.41e-25
Percent Similarity: 36.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6.43%

Length: 422
Matches: 36
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 16

```
US-09-270-437D-5 (1-1708) x US-10-408-765A-2088 (1-422)

Qy   254  GTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    50  ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 69

Qy   314  TCCAAGATAGACGTGCATAGGAAGGAGGAACGCAGGTGCAGCTGAAAAA 361
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    70  SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 85

RESULT 12
US-09-735-705-176
; Sequence 176, Application US/09735705
; Patent No. US2002052329A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Li, Samuel X.
;   APPLICANT: Wang, Aijun
;   APPLICANT: Skeiky, Yasir A.W.
;   APPLICANT: Henderson, Robert A.
;   APPLICANT: McNeill, Patricia D.
;   APPLICANT: Fanger, Neil
;   TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C14
;   CURRENT APPLICATION NUMBER: US/09/735,705
;   CURRENT FILING DATE: 2000-12-12
;   NUMBER OF SEQ ID NOS: 419
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 176
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-735-705-176

Alignment Scores:
Pred. No.:              4.24e-25        Length:         579
Score:                  36.00           Matches:        36
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     9               Gaps:           0

US-09-270-437D-5 (1-1708) x US-09-735-705-176 (1-579)

Qy   254  GTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   207  ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy   314  TCCAAGATAGACGTGCATAGGAAGGAGGAACGCAGGTGCAGCTGAAAAA 361
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   227  SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 13
US-09-735-705-348
; Sequence 348, Application US/09735705
; Patent No. US2002052329A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Li, Samuel X.
;   APPLICANT: Wang, Aijun
;   APPLICANT: Skeiky, Yasir A.W.
;   APPLICANT: Henderson, Robert A.
;   APPLICANT: McNeill, Patricia D.
;   APPLICANT: Fanger, Neil
;   TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C14
;   CURRENT APPLICATION NUMBER: US/09/735,705
;   CURRENT FILING DATE: 2000-12-12
;   NUMBER OF SEQ ID NOS: 419
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-735-705-348

Alignment Scores:
Pred. No.:              4.24e-25        Length:         579
Score:                  36.00           Matches:        36
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     9               Gaps:           0

US-09-270-437D-5 (1-1708) x US-09-735-705-348 (1-579)

Qy   254  GTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   207  ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy   314  TCCAAGATAGACGTGCATAGGAAGGAGGAACGCAGGTGCAGCTGAAAAA 361
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   227  SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 14
US-09-850-716A-176
; Sequence 176, Application US/09850716A
; Patent No. US2002011539A1
; GENERAL INFORMATION:
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: McNeill, Patricia D.
;   APPLICANT: Retter, Marc W.
;   TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C15
;   CURRENT APPLICATION NUMBER: US/09/850,716A
;   CURRENT FILING DATE: 2001-05-07
;   NUMBER OF SEQ ID NOS: 440
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 176
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-850-716A-176

Alignment Scores:
Pred. No.:              4.24e-25        Length:         579
Score:                  36.00           Matches:        36
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            6.43%           Indels:         0
DB:                     9               Gaps:           0

US-09-270-437D-5 (1-1708) x US-09-850-716A-176 (1-579)

Qy   254  GTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   207  ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy   314  TCCAAGATAGACGTGCATAGGAAGGAGGAACGCAGGTGCAGCTGAAAAA 361
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   227  SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 15
```

US-09-850-716A-348
; Sequence 348, Application US/09850716A
; Patent No. US2002011513A1
; GENERAL INFORMATION:
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: McNeill, Patricia D.
;   APPLICANT: Retter, Marc W.
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C15
;   CURRENT APPLICATION NUMBER: US/09/850,716A
;   CURRENT FILING DATE: 2001-05-07
;   NUMBER OF SEQ ID NOS: 440
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-850-716A-348

Alignment Scores:
Pred. No.:   4.24e-25   Length: 579
Score: 36.00   Matches: 36
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 6.43%   Indels: 0
DB: 9   Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-850-716A-348 (1-579)

Qy   254 GTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy   314 TCCAAGATAGACGTGCATAGGAGGAGAACGCAGGTGCAGCTGAAAA 361
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

Search completed: July 13, 2004, 12:49:55
Job time : 94 secs

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:      July 13, 2004, 12:23:53 ; Search time 88 Seconds
                                  (without alignments)
                                  12247.843 Million cell updates/sec

Title:           US-09-270-437D-5
Perfect score:   560
Sequence:        1 agggacgctgccgcaccgcc.........atttccttcaggttttaaaa 1708

Scoring table:   OLIGO
                 Xgapop 60.0 , Xgapext 60.0
                 Ygapop 60.0 , Ygapext 60.0
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        1017041 seqs, 315518202 residues

Word size:       1

Total number of hits satisfying chosen parameters:        2033934

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_13072004_121956_9981/app_query.fasta_1.1863
-DB=SPTREMBL_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN_1_1_146 @runat_13072004_121956_9981 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :    SPTREMBL_25:*
         1:  sp_archea:*
         2:  sp_bacteria:*
         3:  sp_fungi:*
         4:  sp_human:*
         5:  sp_invertebrate:*
         6:  sp_mammal:*
         7:  sp_mhc:*
         8:  sp_organelle:*
         9:  sp_phage:*
        10:  sp_plant:*
        11:  sp_rodent:*
        12:  sp_virus:*
        13:  sp_vertebrate:*
        14:  sp_unclassified:*
        15:  sp_rvirus:*
        16:  sp_bacteriap:*
        17:  sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                            SUMMARIES

Result       %    Query
No.    Score Match Length DB  ID         Description
-----  ----- ----- ------ --- ---------  --------------------------------------
  1     225  40.2    577  11  Q8CGX0     Q8cgx0 rattus norv
```

```
   2  212  37.9  577   4  Q9NZI8   Q9nzi8 homo sapien
   3  166  29.6  577  11  Q8BRH1   Q8brh1 mus musculu
   4  166  29.6  577  11  O88477   O88477 mus musculu
   5  115  20.5  577  11  O80US9   O80us9 mus musculu
   6   77  13.8  576  13  O42254   O42254 gallus gall
   7   41   7.3  594  13  O73932   O73932 xenopus lae
   8   38   6.8  593  13  O57526   O57526 xenopus lae
   9   36   6.4  579   4  O00425   O00425 homo sapien
  10   30   5.4  579  11  Q8C2J9   Q8c2j9 mus musculu
  11   30   5.4  579  11  Q9CPN8   Q9cpn8 mus musculu
  12   24   4.3  582  13  Q9PW80   Q9pw80 brachydanio
  13   19   3.4  556   4  Q9Y6M1   Q9y6m1 homo sapien
  14   16   2.9  169  11  Q7TP50   Q7tp50 rattus norv
  15   15   2.7  545  11  Q7TQF9   Q7tqf9 mus musculu
  16    9   1.6  322  13  Q9PTU0   Q9ptu0 brachydanio
  17    9   1.6  351  16  Q7UXR6   Q7uxr6 rhodopirell
  18    9   1.6  423   4  Q9UIN6   Q9uin6 homo sapien
  19    9   1.6  424   4  Q9UIN7   Q9uin7 homo sapien
  20    9   1.6  424   4  Q9UIP0   Q9uip0 homo sapien
  21    9   1.6  424   4  Q9UBH8   Q9ubh8 homo sapien
  22    9   1.6  434   3  Q872V2   Q872v2 neurospora
  23    9   1.6  499   4  Q9NRH5   Q9nrh5 homo sapien
  24    9   1.6  499   4  Q9NRH4   Q9nrh4 homo sapien
  25    9   1.6  500   4  Q9NPG5   Q9npg5 homo sapien
  26    9   1.6  500   4  Q9NRH7   Q9nrh7 homo sapien
  27    9   1.6  500   4  Q9NPE0   Q9npe0 homo sapien
  28    9   1.6  500   4  Q9NP52   Q9np52 homo sapien
  29    9   1.6  500   4  Q9NRH8   Q9nrh8 homo sapien
  30    9   1.6  500   4  Q9NRH6   Q9nrh6 homo sapien
  31    9   1.6  513   4  Q86Z04   Q86z04 homo sapien
  32    9   1.6  513   7  Q95IF9   Q95if9 homo sapien
  33    9   1.6  528   4  Q8N5P3   Q8n5p3 homo sapien
  34    9   1.6  528   4  Q8IZU6   Q8izu6 homo sapien
  35    9   1.6  528   4  Q8IZU5   Q8izu5 homo sapien
  36    9   1.6  529   4  O43509   O43509 homo sapien
  37    9   1.6  529   4  Q8IZU4   Q8izu4 homo sapien
  38    9   1.6  529   6  Q7YR44   Q7yr44 pan troglod
  39    9   1.6  651  13  Q98SM6   Q98sm6 gallus gall
  40    8   1.4   55   6  O46645   O46645 meles meles
  41    8   1.4   95  10  Q8LLP0   Q8llp0 oryza sativ
  42    8   1.4   95  16  Q97RK6   Q97rk6 streptococc
  43    8   1.4   95  16  Q8CYY6   Q8cyy6 streptococc
  44    8   1.4  133   6  Q9GMU9   Q9gmu9 macaca fasc
  45    8   1.4  136   2  Q93FZ2   Q93fz2 pectobacter
```

```
                              ALIGNMENTS

RESULT 1
Q8CGX0
ID   Q8CGX0;     PRELIMINARY;    PRT;    577 AA.
AC   Q8CGX0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   B-actin zipcode binding protein 1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RA   Eom T., Singer R.H., Bassell G.J.;
RT   "Molecular interactions between rZBP1 and b-actin zipcode required for
RT   transport of mRNA and stimulation of spine growth.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF541940; AAO16210.1; -.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR005004; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
```

```
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
SQ   SEQUENCE   577 AA;   63436 MW;   064767128FBD1EE CRC64;
```

```
Alignment Scores:
Pred. No.:                            577
Score:              2.55e-225         Matches:                225
Percent Similarity:   100.00%         Conservative:             0
Best Local Similarity: 100.00%        Mismatches:               0
Query Match:          40.18%          Indels:                   0
DB:                                   Gaps:                    11
```

```
US-09-270-437D-5 (1-1708) x Q8CGX0 (1-577)

Qy     38  GCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACATC    97
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135  AlaIleMetLysLeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIle   154

Qy     98  CCCGATGAGCAGATAGACACAGGGACCTGAGAATGGGCGCCCAGGGGGCTTTGGCTCTCGG   157
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    155  ProAspGluGlnIleAspThrGlyThrGluAsnGlyArgArgGlyGlyPheGlySerArg   174

Qy    158  GGTCAGCCCCCGCCAGGGCTCACCTGTGGCAGGGCGGGGGCCCAGCAGCAGCAGCAAGTG   217
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    175  GlyGlnProArgGlnGlySerProValAlaLeuAlaGlyAlaProAlaLysGlnGlnGlnVal   194

Qy    218  GACATCCCCTTCGGCCTCGGTGCCCCACCCAGTGGTGGTGCCATTATTGGCAAGGAG    277
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195  AspIleProLeuArgLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGlyGlu   214

Qy    278  GGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAGATAGACGTGCATAGGAAG    337
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    215  GlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLys   234

Qy    338  GAGAAGCAGGTGCAGCTGAAAAAGCCATCAGTGTCACTCCACCCCTGAGGGGCTGCTCC    397
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    235  GluLysGlnValGlnLeuLysLysAlaIleSerValHisSerThrProGluGlyCysSer   254

Qy    398  TCCGCTGTGAAGATGATCTTCGAGATTATGCATAAGAGGCTAAGGACACCAAAACGGCT    457
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255  SerAlaCysLysMetIleLeuGluIleMetHisLysArgLeuArgThrProLysThrAla   274

Qy    458  GACGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTGTGAGGGCGTCTCATTGGCAAG   517
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    275  AspGluValProLeuLysIleLeuAlaHisAsnAsnLeuValGlyArgLeuIleGlyLys   294

Qy    518  GAAGGACGGAACCTGAAGAAGTAGAGAAGATACCGAGACAAAAATACCATCTCCTCG    577
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    295  GluGlyArgAsnLeuLysLysValGluGlnAspThrGluThrLysIleLeuSerSer    314

Qy    578  TTGCAAGACCTTACCCTTACAACCTGAGAGGACCATCACTGTGAGGGGCCATCGAG    637
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315  LeuGlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGlu   334

Qy    638  AATTGTTGCAGGGCCGAGCAGGAAATAATGAAGAAAGTTCGGAGCCATGAGAATGAT    697
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    335  AsnCysCysArgAlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAsp   354

Qy    698  GTGGGCTGCCATGAGC    712
           |||||||||||||||||
Db    355  ValAlaAlaMetSer    359
```

```
RESULT 2
Q9NZI8
ID   Q9NZI8           PRELIMINARY;        PRT;   577 AA.
AC   Q9NZI8;
DT   01-OCT-2000  (TrEMBLrel. 15, Created)
DT   01-OCT-2000  (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003  (TrEMBLrel. 25, Last annotation update)
DE   mRNA-binding protein CRDBP.
```

```
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M.,
RA   Panoutsakopoulos G., Kyriazoglou I., Voutzoulias S., Tsiapalis C.M.,
RA   Kittas C., Agnantis N., Pandis N.;
RT   "Ectopic expression of a KH-domain containing protein, highly
RT   homologous to both human IMP-1 and mouse CRD-BP, in benign and
RT   malignant mesenchymal tumors.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR   EMBL; AF198254; AAF37203.1; -.
HSSP; P11940; 1CVJ.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
SQ   SEQUENCE   577 AA;   63456 MW;   0749A060F252D81D CRC64;
```

```
Alignment Scores:
Pred. No.:                            577
Score:              9.69e-212         Matches:                212
Percent Similarity:   212.00          Conservative:             0
Best Local Similarity: 100.00%        Mismatches:               0
Query Match:          37.86%          Indels:                   0
DB:                                   Gaps:                     4
```

```
US-09-270-437D-5 (1-1708) x Q9NZI8 (1-577)

Qy    725  CCTGGCCTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTCATCCAGCGCAGTCCCGCCG   784
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366  ProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAlaSerSerAlaValProPro    385

Qy    785  CCTCCCAGCAGCGGTTACTGGGGCGTGCTCCCTATAGCTGCTCAGGCTCCCGAGCAG    844
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    386  ProProSerSerValThrGlyValThrGlyAlaAlaProTyrSerSerPheMetGlnAlaProGluGln   405

Qy    845  GAGATGGTGCAGGTGTTATCCCCCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGG    904
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406  GluMetValGlnValPheIleProAlaGlnAlaValGlyAlaIleIleGlyLysLysGly   425

Qy    905  CAGCACATCAAACAGCTCTCCGGTTGCAGCGCCCTCCATCAAGATTGCACCACCCGAA    964
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    426  GlnHisIleLysGlnLeuSerArgPheAlaSerLeuThrIleSerSer    445

Qy    965  ACACCTGACTCCAAAGTTCGTATGGTTATCATCCAAGGAGGAGAGAACTTCTTTGGTCC    1024
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    446  ThrProAspSerLysValArgMetValIleIleThrGlyProGluGlnLysPheHisLys   465

Qy   1025  GCTCAGGGAAGAATCTATGGCAAACTCAAGAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAA   1084
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466  AlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsnPhePheGlyProLysGluGlu    485

Qy   1085  GTGAAGTCTGGAGACCCACCATACGTGTGCCAGCATCAGCAGCTCGGCCGGGTCATTGGCAAA   1144
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486  ValLysLeuGluThrHisIleArgValProAlaSerArgAlaAlaGlyArgValIleGlyLys-Lys   505

Qy   1145  GGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACGGGAGGCAGCTGAGGTGGTAGTACCAAGA   1204
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    506  GlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAlaAlaGluValValProAlaArg   525

Qy   1205  GACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGAGATTCTATGCC   1264
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526  AspGlnThrProAspGluAsnAspGlnValIleValLysIleIleGlyHisPheTyrAla   545
```

```
Qy   1265  AGTCAGATGGCTCAACGGAAGATCCGAGACATCCGGCCCAGGTTAAGCAGCAGCATCAG 1324
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546  SerGlnMetAlaGlnArgLysIleArgAspIleArgLeuLysGlnGlnHisGln 565

Qy   1325  AAGGGACAGAGTAACCAGGCCCAGGCACGGAGGAAG 1360
           |||||||||||||||||||||||||||||||||||
Db    566  LysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 577
```

RESULT 3

```
Q8BRH1
ID   Q8BRH1;       PRELIMINARY;       PRT;   577 AA.
AC   Q8BRH1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Insulin-like growth factor 2.
GN   IGF2BP1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=22354683; PubMed=12466851;
RA   the FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573 (2002).
DR   EMBL; AK044850; BAC32119.1; -.
DR   MGD; MGI:1890357; Igf2bp1.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
SQ   SEQUENCE   577 AA;  63378 MW;  D75897450841532E CRC64;
```

Alignment Scores:

```
Pred. No.:            1.09e-163
Score:               166.00
Percent Similarity:  100.00%
Best Local Similarity: 100.00%
Query Match:          29.64%
DB:                   11
```

```
Length:        577
Matches:       166
Conservative:    0
Mismatches:      0
Indels:          0
Gaps:            0
```

US-09-270-437D-5 (1-1708) x Q8BRH1 (1-577)

```
Qy   215  GTGGACATCCCCCTTCGGCTCTGGTGCCCACCCAGTATGTGGGTGCCCATTATTGGCAAG 274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   194  ValAspIleProLeuArgLeuLeuValProThrGlnTyrValGlyAlaHisTyrTrpGlnGlyLys 213

Qy   275  GAGGGGGCCACCATCCGACAACATCACAAAACAGACCCCAGTCCAAGATAGACCTGCATAGG 334
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   214  GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233

Qy   335  AAGGAGAACCGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCCACTCCACCCTGAGGGCTGC 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   234  LysGluAsnAlaGlyAlaAlaGlyLysProSerValHisSerThrProGluGlyGlyCys 253

Qy   395  TCCTTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACG 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   254  SerSerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThr 273

Qy   455  GCTGACGAGGTTCCCCTGAAGAATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGC 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   274  AlaAspGluValProLeuArgLysIleLeuAlaAlaHisAsnPheValGlyArgLeuIleGly 293

Qy   515  AAGGAAGGACCGAACCTGAAGAAGGTAGACGACCAAGATACCCAGACAAAAATCACCATCTCC 574
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   294  LysGluGlyArgAsnLeuLysLysValGluGlnAspThrGlnThrLysIleThrIleSer 313

Qy   575  TCGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGGCCATC 634
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   314  SerLeuGlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIle 333

Qy   635  GAGAATTGTTGCAGGGCCGAGCAGCAGAAATAATGAAGAAAGTTCGGGAGGCCTATGAGAAT 694
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   334  GluAsnCysCysArgAlaGluGlnGlnLysMetLysLysValAlaArgGluAlaTyrGluAsn 353

Qy   695  GATGTGGCTGCCATGAGC 712
          ||||||||||||||||||
Db   354  AspValAlaAlaMetSer 359
```

RESULT 4

```
Q8477
ID   O8477;        PRELIMINARY;       PRT;   577 AA.
AC   O8477;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Coding region determinant binding protein.
GN   IGF2BP1 OR CRDBP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92217743; PubMed=1559612;
RA   Bernstein P.L., Herrick D.J., Prokipcak R.D., Ross J.;
RT   "Control of c-myc mRNA half-life in vitro by a protein capable of
RT   binding to a coding region stability determinant.";
RL   Genes Dev. 6:642-654(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94158886; PubMed=8114742;
RA   Herrick D.J., Ross J.;
RT   "The half-life of c-myc mRNA in growing and serum-stimulated cells:
RT   influence of the coding and 3' untranslated regions and role of
RT   ribosome translocation.";
RL   Mol. Cell. Biol. 14:2119-2128(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94179348; PubMed=8132663;
RA   Prokipcak R.D., Herrick D.J., Ross J.;
RT   "Purification and properties of a protein that binds to the C-terminal
RT   coding region of human c-myc mRNA.";
RL   J. Biol. Chem. 269:9261-9269(1994).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97322234; PubMed=9178888;
RA   Leeds P., Kren B.T., Boylan J.M., Betz N.A., Steer C.J.,
RA   Gruppuso P.A., Ross J.;
RT   "Developmental regulation of CRD-BP, an RNA-binding protein that
RT   stabilizes c-myc mRNA in vitro.";
RL   Oncogene 14:1279-1286(1997).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Ross J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A.,
RA   Fleisig A.J.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
```

```
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR   EMBL; AF061569; AAC72743.1; -.
DR   EMBL; AK013940; BAB29071.1; -.
DR   HSSP; P11940; 1CVJ.
DR   MGD; MGI:1803357; Igf2bp1.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
SQ   SEQUENCE   577 AA;   63450 MW;   EFBB1AF2F9F0344 CRC64;

Alignment Scores:
Pred. No.:                    1.09e-163       Length:            577
Score:                        166.00          Matches:           166
Percent Similarity:           100.00%         Conservative:      0
Best Local Similarity:        100.00%         Mismatches:        0
Query Match:                  29.64%          Indels:            0
DB:                                           Gaps:              0

US-09-270-437D-5 (1-1708) x 088477 (1-577)
```

Db   456 IleThrGlyProProGluGlnPheLysAlaAlaGlnGlyArgIleTyrGlyLysLeuLys 475

Qy 1055 GAGGAGAACTTCTTTGGTCCCAAGGAGGAGAAGCTGAAGCTGGAGACCCACATACGTGTGCCA 1114
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   476 GluGluAsnPheGlyProLysGluGluLysLeuGluLysLeuGluThrHisIleArgValPro 495

Qy 1115 GCATCAGCAGCTGCCGGTCATTGGCAAAGTGGAAAAACGGTGAACGAGTTGCAGAAT 1174
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   496 AlaSerAlaAlaGlyArgValIleGlyLysLeuGlyLysThrValAsnGluGlnAsn 515

Qy 1175 TTGACGCGCAGCTGAGGTGGTAGTACCAAGAGAACCAGACCCCTGATCGAGAAC 1225
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   516 LeuThrAlaAlaGlyLysThrValAlaAsnGluGlnGlnAsn 531

RESULT 7

ID   O73932   PRELIMINARY;   PRT;   594 AA.
AC   O73932;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   VG1 RNA binding protein variant D.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
RA   Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
RL   Genes Dev. 0:0-0(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98228351; PubMed=9560341;
RA   Deshler J.O., Highett M.I., Abramson T., Schnapp B.J.;
RT   "A highly conserved RNA-binding protein for cytoplasmic mRNA
RT   localization in vertebrates.";
RL   Curr. Biol. 8:489-496(1998).
CC   -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR   EMBL; AF064634; AAC18598.1; -.
DR   EMBL; AF055923; AAC41285.1; -.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
SQ   SEQUENCE 594 AA; 65643 MW; 54CEA7BF0856DD6 CRC64;

Alignment Scores:
Pred. No.:   Length: 594
Score: 4.01e-33   Matches: 41
Percent Similarity: 41.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 7.32%   Gaps: 0
                          13

US-09-270-437D-5 (1-1708) x O73932 (1-594)

Qy   983 CGTATGGTTATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTTCAGGGAAGAATCTAT 1042
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   465 ArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaAlaGlnGlyArgIleTyr 484

Qy 1043 GGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCAC 1102
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   485 GlyLysLeuLysGluGluAsnPheGlyProLysGluGluValLysLeuGluThrHis 504

Qy 1103 ATA 1105
          |||
Db   505 Ile 505

RESULT 8

Db   456 IleThrGlyProProGluGlnPheLysAlaAlaGlnGlyArgIleTyrGlyLysLeuLys 475

Qy 1055 GAGGAGAACTTCTTTGGTCCCAAGGAGGAGAAGCTGAAGCTGGAGACCCACATACGTGTGCCA 1114
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   476 GluGluAsnPheGlyProLysGluGluLysLeuGluLysLeuGluThrHisIleArgValPro 495

Qy 1115 GCATCAGCAGCTGCCGGTCATTGGCAAAGTGGAAAAACGGTGAAAAACGGTGAACGAGTTGCAGAAT 1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   496 AlaSerAlaAlaGlyArgValIleGlyLysLeuGlyLysThrValAsnGluGlnGlnAsn 515

Qy 1175 TTGACGCGCAGCTGAGGTGGTAGTACCAAGAGAACCAGACCCCTGATCGAGAAC 1234
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   516 LeuThrArgAlaAlaGluGluValValProArgArgAspAspProAspArgGlnGlnAsn 535

Qy 1235 ATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCAGAC 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   536 IleValLysIleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAsp 555

Qy 1295 ATCCTGGCCCAGGTTAAGCAGCAGCATCAGAGAGGACAGAGTAAC 1339
          |||||||||||||||||||||||||||||||||||||||||||||
Db   556 IleLeuAlaGlnValLysGlnGlnHisGlnArgGlyGlnSerAsn 570

RESULT 6

ID   O42254   PRELIMINARY;   PRT;   576 AA.
AC   O42254;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Zipcode-binding protein.
GN   ZBP1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97220007; PubMed=9121465;
RA   Ross A.F., Oleynikov Y.S., Kislauskis E.H., Taneja K.L., Singer R.H.;
RT   "Characterization of a beta-actin mRNA zipcode-binding protein.";
RL   Mol. Cell. Biol. 17:2158-2165(1997).
CC   -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR   EMBL; AF026527; AAB82295.1; -.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
SQ   SEQUENCE 576 AA; 63271 MW; 01AAF2D1D81C8811 CRC64;

Alignment Scores:
Pred. No.:   Length: 576
Score: 1e-70   Matches: 77
Percent Similarity: 77.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 13.75%   Gaps: 0
                          13

US-09-270-437D-5 (1-1708) x O42254 (1-576)

Qy   995 ATCACTGGACCGCCAGAGGCCCAATTCAAGGCTTCAGGGAAGAATCTATGGCAAACTCAAG 1054
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   455 IleThrGlyProProGluAlaGlnPheLysAlaAlaGlnGlyArgIleTyrGlyLysLeuLys 474

Qy 1055 GAGGAGAACTTCTTTGGTCCCAAGGAGGAGAAGCTGAAGCTGGAGACCCACATACGTGTGCCA 1114
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
O57526
ID   O57526       PRELIMINARY;      PRT;   593 AA.
AC   O57526;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   KH domain-containing transcription factor B3.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92249652; PubMed=1577195;
RA   Pfaff S.L., Taylor W.L.;
RT   "Characterization of a Xenopus oocyte factor that binds to a
RT   developmentally regulated cis-element in the TFIIIA gene.";
RL   Dev. Biol. 151:306-316(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Griffin D., Taylor W.L.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
RA   Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
RL   Genes Dev. 0:0-0(1998).
CC   -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR   EMBL; AF042353; AAB97457.1; -.
DR   EMBL; AF064633; AAC18597.1; -.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
SQ   SEQUENCE  593 AA;  65385 MW;  5A5AB4B4A1D55DF7 CRC64;

Alignment Scores:
Pred. No.:                       5.46e-30      Length:         593
Score:                           38.00         Matches:        38
Percent Similarity:              100.00%       Conservative:   0
Best Local Similarity:           100.00%       Mismatches:     0
Query Match:                     6.79%         Indels:         0
DB:                                            Gaps:           13

US-09-270-437D-5 (1-1708) x O57526 (1-593)

Qy    983 CGTATGGTTATCATCACTGGACGCCGCCAATTCAAGGCTCAGGGAAGAATCTAT 1042
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    464 ArgMetValIleIleThrGlyArgProProGluAlaGlnPheLysAlaGlnGlyArgIleTyr 483

Qy   1043 GGCAAACTCAAGGAGGAGAACTTCTTTGGTCCAAGGAGGAAGAGTGAAGTGGAG 1096
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 GlyLysLeuLysGluGluAsnPhePheGlyProLysArgLysSerGluValGlu 501

RESULT 9
O00425
ID   O00425       PRELIMINARY;      PRT;   579 AA.
AC   O00425;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative RNA binding protein KOC (KOC).
GN   KOC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mueller-Pillasch F., Lacher U., Wallrapp C., Et AL.;
RL   Oncogene 0:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RA   Mueller-Pillasch F., Lacher U., Wallrapp C.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR   EMBL; U97188; AAC35208.1; -.
DR   EMBL; U76705; AAD09223.1; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0003723; F:RNA binding; TAS.
DR   GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR   GO; GO:0006412; P:protein biosynthesis; TAS.
DR   GO; GO:0006396; P:RNA processing; TAS.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
DR   PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
SQ   SEQUENCE  579 AA;  63720 MW;  AE5C3A8EE3C135C5 CRC64;

Alignment Scores:
Pred. No.:                       6.72e-28      Length:         579
Score:                           36.00         Matches:        36
Percent Similarity:              100.00%       Conservative:   0
Best Local Similarity:           100.00%       Mismatches:     0
Query Match:                     6.43%         Indels:         0
DB:                                            Gaps:           0

US-09-270-437D-5 (1-1708) x O00425 (1-579)

Qy    254 GTGGGTGCCATTATTGGCAAGGAGGAGGGGGCCAACATCCGCAAACTACAAAAACAGACCCAG 313
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 ValGlyAlaIleIleGlyLysGluGluGlyAlaAsnIleArgAsnIleThrLysGlnThrGln 226

Qy    314 TCCAAGATAGACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAAAAA 361
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 10
Q8C2J9
ID   Q8C2J9       PRELIMINARY;      PRT;   579 AA.
AC   Q8C2J9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Insulin-like growth factor 2.
GN   IGF2BP3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NOD; TISSUE=Thymus;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK088465; BAC40370.1; -.
DR   MGD; MGI:1890359; Igf2bp3.
```

DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
SQ   SEQUENCE 579 AA; 63551 MW; 937E601A95D06B77 CRC64;

Alignment Scores:
Pred. No.:   1.24e-21   Length: 579
Score: 30.00   Matches: 30
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 5.36%   Indels: 0
DB: 11   Gaps: 0

US-09-270-437D-5 (1-1708) x Q8CJ9 (1-579)

Qy   254 GTGGGTGCCATTATTGGCAAGGAGGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   207 ValGlyAlaIleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy   314 TCCAAGATAGACGTGCATAGGAAGGAGGAGAAC 343
           |||||||||||||||||||||||||||||||||
Db   227 SerLysIleAspValHisArgLysGluAsn 236

RESULT 11
Q9CPN8
ID   Q9CPN8   PRELIMINARY;   PRT;   579 AA.
AC   Q9CPN8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   10 days embryo cDNA, RIKEN full-length enriched library,
DE   clone:2610036B18, full insert sequence (Igf2 mRNA-binding protein 3)
DE   (Insulin-like growth factor 2, binding protein 3).
GN   IGF2BP3 OR 2610101N1RIK OR MIMP3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,
RA   Yuasa Y., Takeda M., Okano H.;

RT   "Expression of mouse igf2 mRNA-binding protein 3 and its implications
RT   for the developing central nervous system.";
RL   J. Neurosci. Res. 0:0-0(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
RA   Strausberg R.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR   EMBL; AK011689; BAB27779.1; -.
DR   EMBL; AB046173; BAB19755.1; -.
DR   EMBL; BC045138; AAH45138.1; -.
DR   EMBL; BC049082; AAH49082.1; -.
DR   MGD; MGI:1890359; Igf2bp3.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
SQ   SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

Alignment Scores:
Pred. No.:   1.24e-21   Length: 579
Score: 30.00   Matches: 30
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 5.36%   Indels: 0
DB: 11   Gaps: 0

US-09-270-437D-5 (1-1708) x Q9CPN8 (1-579)

Qy   254 GTGGGTGCCATTATTGGCAAGGAGGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   207 ValGlyAlaIleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy   314 TCCAAGATAGACGTGCATAGGAAGGAGGAGAAC 343
           |||||||||||||||||||||||||||||||||
Db   227 SerLysIleAspValHisArgLysGluAsn 236

RESULT 12
Q9PW80
ID   Q9PW80   PRELIMINARY;   PRT;   582 AA.
AC   Q9PW80;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)

```
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Vg1 RNA binding protein (Decapentaplegic and Vg-related 1, RNA binding
DE   protein).
GN   DVR1RBP.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhang Q., Yaniv K., Oberman F., Wolke U., Git A., Fromer M.,
RA   Taylor W., Meyer D., Standart N., Raz E., Yisraeli J.K.;
RT   "Vg1 RBP intracellular distribution and evolutionarily conserved
RT   expression suggest multiple roles during development.";
RL   Mech. Dev. 0:0-0(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Body;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR   EMBL; AF161270; AAD45610.1; -.
DR   EMBL; BC045873; AAH45873.1; -.
DR   ZFIN; ZDB-GENE-000308-1; dvr1rbp.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
SQ   SEQUENCE 582 AA; 63351 MW; 9DAE63200681B306 CRC64;

Alignment Scores:
Pred. No.:                           2.3e-15
Score:                               24.00
Percent Similarity:                  100.00%
Best Local Similarity:               100.00%
Query Match:                         4.29%
DB:                                  13

                                     Length:            582
                                     Matches:           24
                                     Conservative:      0
                                     Mismatches:        0
                                     Indels:            0
                                     Gaps:              0

US-09-270-437D-5 (1-1708) x Q9PW80 (1-582)

Qy   467 CCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAGGAGGACGG 526
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   277 ProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArg 296
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   527 AACCTGAAGAG 538
         |||||||||||
Db   297 AsnLeuLysLys 300

RESULT 13
Q9Y6M1
ID   Q9Y6M1       PRELIMINARY;      PRT;    556 AA.
AC   Q9Y6M1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hepatocellular carcinoma autoantigen.
GN   P62.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9207072; PubMed=10190901;
RA   Zhang J.Y., Chan E.K., Peng X.X., Tan E.M.;
RT   "A novel cytoplasmic protein with RNA-binding motifs is an autoantigen
RT   in human hepatocellular carcinoma.";
RL   J. Exp. Med. 189:1101-1110(1999).
CC   -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR   EMBL; AF057352; AAD31596.1; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0003723; F:RNA binding; TAS.
DR   GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR   GO; GO:0009386; P:translational attenuation; TAS.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   InterPro; IPR000504; RNA_rec_mot.
DR   Pfam; PF00013; KH; 4.
DR   Pfam; PF00076; rrm; 2.
DR   SMART; SM00322; KH; 4.
DR   SMART; SM00360; RRM; 2.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
DR   PROSITE; PS50102; RRM; 2.
SQ   SEQUENCE 556 AA; 61842 MW; 1EDEFB10044 3DDC4 CRC64;

Alignment Scores:
Pred. No.:                           3.87e-10
Score:                               19.00
Percent Similarity:                  100.00%
Best Local Similarity:               100.00%
Query Match:                         3.39%
DB:                                  4

                                     Length:            556
                                     Matches:           19
                                     Conservative:      0
                                     Mismatches:        0
                                     Indels:            0
                                     Gaps:              0

US-09-270-437D-5 (1-1708) x Q9Y6M1 (1-556)

Qy   983 CGTATGGTTATCATCACTGGACCCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATC 1039
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   431 ArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 449

RESULT 14
Q7TP50
ID   Q7TP50       PRELIMINARY;      PRT;    169 AA.
AC   Q7TP50;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Ab2-255.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,
```

```
RA    Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA    Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RT    "Liver regeneration after PH.";
RL    Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR    EMBL; AY325199; AAP92600.1; -.
SQ    SEQUENCE    169 AA;   17743 MW;  19BF6295C1000CBA CRC64;

Alignment Scores:
Pred. No.:              6.17e-07          Length:              169
Score:                  16.00             Matches:             16
Percent Similarity:     100.00%           Conservative:        0
Best Local Similarity:  100.00%           Mismatches:          0
Query Match:            2.86%             Indels:              0
DB:                     11                Gaps:                0

US-09-270-437D-5 (1-1708) x Q7TP50 (1-169)

Qy   977 AAAGTTCCTATGGTTATCATCACTGGACCGCCAGAGGCCCAATTCAAG 1024
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db    59 LysValArgMetValIleIleThrGlyProProGluAlaGlnPheLys 74


RESULT 15
Q7TQF9
ID   Q7TQF9          PRELIMINARY;          PRT;    545 AA.
AC   Q7TQF9;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Mus musculus (Mouse)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC054552; AAH54452.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE    545 AA;   59679 MW;  88DD35FF30DF3091 CRC64;

Alignment Scores:
Pred. No.:              5.85e-06          Length:              545
Score:                  15.00             Matches:             15
Percent Similarity:     100.00%           Conservative:        0
Best Local Similarity:  100.00%           Mismatches:          0
Query Match:            2.68%             Indels:              0
DB:                     11                Gaps:                0
```

US-09-270-437D-5 (1-1708) x Q7TQF9 (1-545)

```
Qy   494 TTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAG 538
         |||||||||||||||||||||||||||||||||||||||||||||
Db   277 PheValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLys 291
```

Search completed: July 13, 2004, 12:34:38
Job time : 100 secs

```
            GenCore version 5.1.6
       Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    July 13, 2004, 23:27:48 ; Search time 134 Seconds
                                     (without alignments)
                                     7073.553 Million cell updates/sec

Title:         US-09-270-437D-5
Perfect score: 1708
Sequence:      1 aggaggctgccgcaccgcc.......atttccttcagtttaaaa 1708

Scoring table: OLIGO_NUC
               Gapop 60.0 , Gapext 60.0

Searched:      682709 seqs, 277475446 residues

Word size :    0

Total number of hits satisfying chosen parameters:    1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   Issued Patents NA:*
        1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
        2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
        3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
        4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
        5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
        6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1708	100.0	1708	3	US-09-061-709-5	Sequence 5, Appli
2	1708	100.0	1708	4	US-09-899-651-5	Sequence 5, Appli
3	1672	97.9	1946	4	US-09-061-709-7	Sequence 7, Appli
4	1672	97.9	1946	4	US-09-899-651-7	Sequence 7, Appli
5	31	1.8	2224	3	US-09-261-855-1	Sequence 1, Appli
c 6	24	1.4	24	3	US-09-261-855-38	Sequence 38, Appl
7	23	1.3	23	3	US-09-261-855-36	Sequence 36, Appl
8	20	1.2	1740	4	US-09-643-597-347	Sequence 347, App
9	20	1.2	1740	4	US-09-542-615A-347	Sequence 347, App
10	20	1.2	1740	4	US-09-605-421B-347	Sequence 347, App
11	20	1.2	4159	4	US-09-061-709-4	Sequence 4, Appli
12	20	1.2	4159	4	US-09-899-651-4	Sequence 4, Appli
13	20	1.2	4181	4	US-09-643-597-175	Sequence 175, App
14	20	1.2	4181	4	US-09-480-884A-175	Sequence 175, App
15	20	1.2	4181	4	US-09-542-615A-175	Sequence 175, App
16	20	1.2	4181	4	US-09-606-421B-175	Sequence 175, App
17	18	1.1	324	4	US-09-489-039A-4791	Sequence 4791, Ap
c 18	18	1.1	513	4	US-09-252-991A-7364	Sequence 7364, Ap
19	18	1.1	1001	4	US-09-671-317-286	Sequence 286, App
20	18	1.1	1059	4	US-09-252-991A-7254	Sequence 7254, Ap
21	18	1.1	1233	4	US-09-489-039A-4691	Sequence 4691, Ap
22	18	1.1	1408	1	US-08-447-554-3	Sequence 3, Appli
23	18	1.1	1408	1	US-08-448-160-3	Sequence 3, Appli
c 24	18	1.1	3263	3	US-08-714-918-20	Sequence 20, Appl
c 25	18	1.1	3263	3	US-09-265-315-20	Sequence 20, Appl
c 26	18	1.1	3263	3	US-09-265-315-20	Sequence 20, Appl
c 27	18	1.1	3263	3	US-09-266-417-20	Sequence 20, Appl
c 28	18	1.1	3263	4	US-09-528-709-20	Sequence 20, Appl
c 29	18	1.1	3263	4	US-09-527-745-20	Sequence 20, Appl
c 30	18	1.1	4165	1	US-08-095-737-1	Sequence 1, Appli
c 31	18	1.1	4165	1	US-08-480-145-1	Sequence 1, Appli
c 32	18	1.1	4165	2	US-08-477-389-1	Sequence 1, Appli
c 33	18	1.1	5191	5	US-08-340-428B-1	Sequence 1, Appli
c 34	18	1.1	5191	5	PCT-US93-07306-1	Sequence 1, Appli
c 35	18	1.1	5333	6	US-08-956-171E-170	Sequence 170, App
c 36	18	1.1	8224	6	5180808-1	Patent No. 5180808
c 37	18	1.1	9785	1	US-09-479-128-1	Sequence 1, Appli
c 38	18	1.1	786431	4	US-09-751-389-3	Sequence 3, Appli
39	18	1.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
40	18	1.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 41	17	1.0	17	3	US-09-261-855-33	Sequence 33, Appl
c 42	17	1.0	47	3	US-09-422-978-2996	Sequence 2996, Ap
43	17	1.0	189	5	US-09-511-625B-34	Sequence 34, Appl
c 44	17	1.0	312	4	US-09-489-039A-6421	Sequence 6421, Ap
c 45	17	1.0	462	4	US-09-134-000C-3161	Sequence 3161, Ap

ALIGNMENTS

```
RESULT 1
US-09-061-709-5
; Sequence 5, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
;  APPLICANT: Chen, Yao-Tseng
;  APPLICANT: Gure, Ali
;  APPLICANT: Tsang, Solam
;  APPLICANT: Stockert, Elisabeth
;  APPLICANT: Jager, Elke
;  APPLICANT: Knuth, Alexander
;  APPLICANT: Old, Lloyd J.
;  TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
;  TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
;  FILE REFERENCE: LUD 5538
;  CURRENT APPLICATION NUMBER: US/09/061,709B
;  CURRENT FILING DATE: 1998-04-17
;  NUMBER OF SEQ ID NOS: 8
SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-5

  Query Match             100.0%;  Score 1708;  DB 3;  Length 1708;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1708;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 AGGGACGGCTGCCGCACCGCCCCAGTTACCCGGGAGCCATCATGAAGCTGAATGGCCA 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGGGACGGCTGCCGCACCGCCCCAGTTACCCGGGAGCCATCATGAAGCTGAATGGCCA 60

Qy     61 CCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACATCCCGATGAGCAGATAGCACAGGG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACATCCCGATGAGCAGATAGCACAGGG 120

Qy    121 ACCTGAGAATGGGCGCGAGGGGGCTTTGGCTCTCGGGGTCAGCCCCAGGGCTCACC 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACCTGAGAATGGGCGCGAGGGGGCTTTGGCTCTCGGGGTCAGCCCCAGGGCTCACC 180

Qy    181 TGTGGGCAGCGGGGGCCCCAGCCAGCAGCAAGTGGACATCCCCTTCGGCTCCTGT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTGGGCAGCGGGGGCCCCAGCCAGCAGCAAGTGGACATCCCCTTCGGCTCCTGT 240

Qy    241 GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCAC 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCAC 300

Qy    301 AAAACAGACCCAGTCCAAGATAGACCTGCATAGGAACGAGGGAGAACGCAGGTGCAGCTGAAAA 360
```

```
Qy    301  AAAACAGACCCAGTCCAAGATAGACGTGCATAGGAACGGAGGAGACGCAGTGCAGCTGAAAA  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db

Qy    361  AGCCATCAGTGTGCACTCCACCCTGAGGGCTGTCGCTCCTCCCGCGTTCCCCGCCCCTGAGGATCTTGGA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AGCCATCAGTGTGCACTCCACCCTGAGGGCTGTCGCTCCTCCCGCGTTGTAAGATGATCTTGGA  420

Qy    421  GATTATGCATAAAGAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCCTGAAGATCCT  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GATTATGCATAAAGAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCCTGAAGATCCT  480

Qy    481  GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGAGGACGGAACGTGAAGAAGT  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGAGGACGGAACGTGAAGAAGT  540

Qy    541  AGAGCAAGATACCGAGACAAAAATCACCATCTCGTTGCAGAATTGTTGCAGGCCGAGCAGGA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  AGAGCAAGATACCGAGACAAAAATCACCATCTCGTTGCAAGACCTTACCCTTTACAA  600

Qy    601  CCCTGAGGAGGACCATCACTGTGAAGGGGCCATCGAGAATGTGCTGCCATGAGCTCTCCACCT  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CCCTGAGGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGCCGAGCAGGA  660

Qy    661  AATAATGAGGAAGAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCAGCTTCATCCAGCGCAGTCCCT  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AATAATGAAGAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCAGCTTCATCCAGCGCAGTCCACCT  720

Qy    721  GATCCCCTGGCCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GATCCCCGGCCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCC  780

Qy    781  GCCGCCTTCCCAGCAGCGTTACTGGGCTGCTTCCCTATAGCTTCCTTTATGCAGGCTCCCGA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GCCGCCTTCCCAGCAGCGTTACTGGGCTGCTCCCTATAGCTTCCTTTATGCAGGCTCCCGA  840

Qy    841  GCAGGAGATGGTGCAGGGTGTTATCCCGCCAGTGGCGCCATCATCGGCAAGAA  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  GCAGGAGATGGTGCAGGGTGTTATCCCGCCAGTGGCGCCATCATCGGCAAGAA  900

Qy    901  GGGGCAGCACATCAAACAGCTCTCCCGGTTTGCCAGCGCGCCTCCATCAAGATTGCACCACC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  GGGGCAGCACACATCAAACAGCTCTCCCGGTTTGCCAGCGCGCCTCCATCAAGATTGCACCACC  960

Qy    961  CGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTCGACCGCCAGAGGCCCAATT  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  CGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGCCCAATT  1020

Qy   1021  CAAGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGTCCCAAGGA  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  CAAGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGTCCCAAGGA  1080

Qy   1081  GGAAGTGAAGCTGGAGACCCACCATACGTGTGCCAGCATCAGCAGTGGCCCGGGTCATTGG  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  GGAAGTGAAGCTGGAGACCCACCATACGTGTGCCAGCATCAGCAGTGGCCCGGGTCATTGG  1140
                                                                    Gaps  0:

Qy   1141  CAAAGGTGGAAAAACGGTGAACGAGTTGACGGCAGTGGTAGTACC  1200
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CAAAGGTGGAAAAACGGTGAACGAGTTGACGGCAGTGGTAGTACC  1200

Qy   1201  AAGAGACCAGACCCCTGATGAGAACGAACCAGTCATCGTTGAAAATCATCGGACATTTCTA  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  AAGAGACCAGACCCCTGATGAGAACGAACCAGTCATCGTGAAAATCATCGGACATTTCTA  1260

Qy   1261  TGCCAGTCAGATGGCTCAACGGTCCAACGGATCCTGGACTCCTGGCCCAGTTAAGCAGCAGCA  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  TGCCAGTCAGATGGCTCAACGGCTCAACGGATCCTGGACGCACGGAAGTGACCAGCCCCTGTCC  1320

Qy   1321  TCAGAGGACAGAGTAACCAGGCCCCAGGCGGAGGAAGTGACATCCCCTGTCC  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  TCAGAGGACAGAGTAACCAGGCCCCAGGCGGAGGAAGTGACATCCCCTGTCC  1380

Qy   1381  CTTNGAGTCCAGGACAACAACGGGCAGAGAAATCGAGAGAGTGTGCTCCCCGGCGAGCCTCGA  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db
```

```
Db   1381  CTTNGAGTCCAGGACAACAACGGGCAGAGAAATCGAGAGAGTGTGCTCCCCGGCGAGCCTCGA  1440

Qy   1441  GAATGAGTGGGAATCGGGGACACNTGGGCGGGGCTGTGAGATCAGGTTTGCCCACTTGATT  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  GAATGAGTGGGAATCGGGGACACNTGGGCGGGGCTGTAGATCAGGTTTGCCCACTTGATT  1500

Qy   1501  GAGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCAAACACCCAATTGGC  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  GAGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCAAACACCCAATTGGC  1560

Qy   1561  CCAACACTGTNTGCCCTCGGGGTGTCAGAAATTNTAGCGCAAGGCACTTTAAACGTGG  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  CCAACACTGTNTGCCCTCGGGGTGTCAGAAATTNTAGCGCAAGGCACTTTAAACGTGG  1620

Qy   1621  ATTGTTTAAAGAAGCTCTCCAGGCCCCAAGAGGGTGGATCACACCTCAGTGGGAAGA  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  ATTGTTTAAAGAAGCTCTCCAGGCCCCAAGAGGGTGGATCACACCTCAGTGGGAAGA  1680

Qy   1681  AAAATAAAAATTTCCTTCAGTTTTAAAA  1708
           ||||||||||||||||||||||||||||
Db   1681  AAAATAAAAATTTCCTTCAGTTTTAAAA  1708
```

```
Query Match       100.0%;  Score 1708;   DB 4;   Length 1708;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1708;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1  AGGGACGCGTCGCCACGCCCCAGTTTACCCCGGGAGGCCATCATGAGAGCTGAATGGCCA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  AGGGACGCGTCGCCACGCCCCAGTTTACCCCGGGAGGCCATCATGAGAGCTGAATGGCCA   60

Qy     61  CCAGTTGGAGAACCACCATGCCCTGAAGGTCCTACATCCCGATGAGCAGATAGCACAGGG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CCAGTTGGAGAACCACCATGCCCTGAAGGTCCTACATCCCGATGAGCAGATAGCACAGGG  120

Qy    121  ACCTGAGAATGGGCGCCTTTGGCTCTCGGGGGGCTTTGGCTCGCAAGCAGCCCCCAGGCTCACC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  ACCTGAGAATGGGCGCCTTTGGCTCTCGGGGGGCTTTGGCTCGCAAGCAGCCCCCAGGCTCACC  180

Qy    181  TGTGGCAGCGGGGGCGCCCCAGCCCAGCAGTGACATGACATCCCCTTCGGCTCCTGT  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TGTGGCAGCGGGGGCGCCCCAGCCCAGCAGTGACATGACATCCCCTTCGGCTCCTGT  240

Qy    241  GCCCACCCAGTATGGTGGGTGCCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCAC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db
```

Qy   1381 CTTNGAGTGTCCAGGACAACAACCGGCCAGCAGAAATCGAGAGTGTGCTCTCCCGGCAGGCCTGA 1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CTTNGAGTGTCCAGGACAACAACCGGCCAGCAGAAATCGAGAGTGTGCTCTCCCGGCAGGCCTGA 1440

Qy   1441 GAATGAGTGGGAATCCGGGACACNTGGGCCGGGCTGTAGATCAGGTTGCCCACTGATT 1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GAATGAGTGGGAATCCGGGACACNTGGGCCGGGCTGTAGATCAGGTTGCCCACTGATT 1500

Qy   1501 GAGAAAGATGTTCCAGTGAGGAACCCTGATCNTNCAGCCCCAAACACCCACCAATTGGC 1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAGAAAGATGTTCCAGTGAGGAACCCTGATCNTNCAGCCCCAAACACCCACCAATTGGC 1560

Qy   1561 CCAACACTGTNTGCCCCTCGGGGTGTCAGAAATTNTAGCGCAAGGCACTTTTTAAACGTGG 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCAACACTGTNTGCCCCTCGGGGTGTCAGAAATTNTAGCGCAAGGCACTTTTTAAACGTGG 1620

Qy   1621 ATTGTTTAAAGAAGCTCTCCAGGCCCCACCAAGAGAGGTGGATCACACCTCAGTGGGAAGA 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ATTGTTTAAAGAAGCTCTCCAGGCCCCACCAAGAGAGGTGGATCACACCTCAGTGGGAAGA 1680

Qy   1681 AAAATAAAATTTCCTTCAGGTTTTAAAA 1708
           ||||||||||||||||||||||||||||
Db   1681 AAAATAAAATTTCCTTCAGGTTTTAAAA 1708

RESULT 3
US-09-061-709-7
; Sequence 7, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
;   APPLICANT: Chen, Yao-Tseng
;   APPLICANT: Gure, Ali
;   APPLICANT: Tsang, Solam
;   APPLICANT: Stockert, Elisabeth
;   APPLICANT: Jager, Elke
;   APPLICANT: Knuth, Alexander
;   APPLICANT: Old, Lloyd J.
;   TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
;   TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
;   FILE REFERENCE: LUD 5538
;   CURRENT APPLICATION NUMBER: US/09/061,709B
;   CURRENT FILING DATE: 1998-04-17
;   NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
;   LENGTH: 1946
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-061-709-7

Query Match   97.9%; Score 1672; DB 3; Length 1946;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   37 AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT 96
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   275 AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT 334

Qy   97 CCCCGATGAGCAGATAGCACAGGGACCTGAGGTGGTAGTACC 156
          ||||||||||||||||||||||||||||||||||||||||||
Db   335 CCCCGATGAGCAGATAGCACAGGGACCTGAGGTGGTAGTACC 394

Qy   157 GGGTCAGCCCCGCCAGGGCTCACCTGTGGCCCAGCAGCAGCAAGT 216
          |||||||||||||||||||||||||||||||||||||||||||||
Db   395 GGGTCAGCCCCGCCAGGGCTCACCTGTGGCCCAGCAGCAGCAAGT 454

Qy   217 GGACATCCCCTTCGGCTCCGGTGCCCACCAGTATGTGGTGCCATTATTGGCAAGGA 276
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   455 GGACATCCCCTTCGGCTCCGGTGCCCACCAGTATGTGGTGCCATTATTGGCAAGGA 514

Qy   277 GGGGGCCACCATCCCGCAACATCACAAAACAGACCCAAGATAGAGCGTGCATAGGAA 336

Db   241 GCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCAC 300

Qy   301 AAAACAGACCCAGTCCAAGATAGAGCGTGCATAGAGACGTGCAGCAGCGAGCTGAAAA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AAAACAGACCCAGTCCAAGATAGAGCGTGCATAGAGACGTGCAGCAGCGAGCTGAAAA 360

Qy   361 AGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCCGCTTGTAAGATGATCTTGGA 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 AGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCCGCTTGTAAGATGATCTTGGA 420

Qy   421 GATTATGCATAAAGAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTGAAGATCCT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GATTATGCATAAAGAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTGAAGATCCT 480

Qy   481 GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAGGT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAGGT 540

Qy   541 AGAGCAAGATACCGAGACAAAAATCACCATCACCATCTCCGTTGCAAGACCTTACCCTTACAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AGAGCAAGATACCGAGACAAAAATCACCATCACCATCTCCGTTGCAAGACCTTACCCTTACAA 600

Qy   601 CCCTGAGAGGACCATCACTGTGAAGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGCAGGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CCCTGAGAGGACCATCACTGTGAAGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGCAGGA 660

Qy   661 AATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCT 720

Qy   721 GATCCCTGGCCTGAACCTGGCTGCTGCTAGTGCTTCTTTTCCCAGCTTCATCCAGCGCAGTCCCC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GATCCCTGGCCTGAACCTGGCTGCTGCTAGTGCTTCTTTTCCCAGCTTCATCCAGCGCAGTCCCC 780

Qy   781 GCCGCCCTCCAGCAGCGGTTACTGGGCTGCTCCTATAGCTCCTTTATGCAGGCTCCCGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GCCGCCCTCCAGCAGCGGTTACTGGGCTGCTCCTATAGCTCCTTTATGCAGGCTCCCGA 840

Qy   841 GCAGGAGATGGTGCAGTGTTATCCCGCGGCAGTGGGCGCCATCATCGGCAAGAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GCAGGAGATGGTGCAGTGTTATCCCGCGGCAGTGGGCGCCATCATCGGCAAGAA 900

Qy   901 GGGGCAGCACATCAAAACAGCTCTCCCGGTTTGCCAGCGCCTCATCAAGATTGCACCACC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GGGGCAGCACATCAAAACAGCTCTCCCGGTTTGCCAGCGCCTCATCAAGATTGCACCACC 960

Qy   961 CGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCCGCCAGAGGCCCAATT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCCGCCAGAGGCCCAATT 1020

Qy   1021 CAAGGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTTGGTCCCAAGGA 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAAGGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTTGGTCCCAAGGA 1080

Qy   1081 GGAAGTGAAGCTCGGAGAACCATACGTGTGCCAGCATCAGCAGTTCATTGG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGAAGTGAAGCTCGGAGAACCATACGTGTGCCAGCATCAGCAGTTCATTGG 1140

Qy   1141 CAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTACC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTACC 1200

Qy   1201 AAGAGACCAGCCCCTGATGAGCAGATAGCACAGGGTCATCGTGAAAATCATCCGGACATTTCTA 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAGAGACCAGCCCCTGATGAGCAGATAGCACAGGGTCATCGTGAAAATCATCCGGACATTTCTA 1260

Qy   1261 TGCCAGTGCAGTGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAGCAGCA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGCCAGTGCAGTGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAGCAGCA 1320

Qy   1321 TCAGAAGGGACAGAGTAACCAGGCCCAGGCACGGCAGGAAGTGACCAGCCCCCTGTCC 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TCAGAAGGGACAGAGTAACCAGGCCCAGGCACGGCAGGAAGTGACCAGCCCCCTGTCC 1380

Db 515 GGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAA 574
Qy 337 GGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTC 396
Db 575 GGAGAACCCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTC 634
Qy 397 CTCCGCTTGTAGAGATGATCTTGGAGATTATGCATAAAGAGGTAAGGACACCAAAACGGC 456
Db 635 CTCCGCTTGTAGAGATGATCTTGGAGATTATGCATAAAGAGGTAAGGACACCAAAACGGC 694
Qy 457 TGACGAGGTTCCCCTGAAGATCCTGCCCATAATAACTTGTGAGGCGTCTCATTGGCAA 516
Db 695 TGACGAGGTTCCCCTGAAGATCCTGCCCATAATAACTTGTGAGGCGTCTCATTGGCAA 754
Qy 517 GGAAGGACGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTC 576
Db 755 GGAAGGACGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTC 814
Qy 577 GTTGCAAGACCTTACCCTTACACACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA 636
Db 815 GTTGCAAGACCTTACCCTTACACACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA 874
Qy 637 GAATTGTTGCAGGGCCGAGCAGGAGAAATAATGAAGAAGTTCGGGAGGCCTATGAGAATGA 696
Db 875 GAATTGTTGCAGGGCCGAGCAGGAGAAATAATGAAGAAGTTCGGGAGGCCTATGAGAATGA 934
Qy 697 TGTGGCTGCCATGAGCAGCTTCACCTGGCCTTCACCTGCCTGCACCTGGCTGTGGTCTTTT 756
Db 935 TGTGGCTGCCATGAGCAGCTTCACCTGGCCTTCACCTGCCTGCACCTGGCTGTGGTCTTTT 994
Qy 757 CCCAGCTTCATCAGCGGCAGTCCGCCGCCCTCCCAGCGGCGTTACTCGGGCGTGCTCCCTA 816
Db 995 CCCAGCTTCATCAGCGGCAGTCCGCCGCCCTCCCAGCGGCGTTACTGGGCGTGCTCCCTA 1054
Qy 817 TAGCTCCTTTATGCAGGCTCCCCGAGCAGCAGGAGATGGTGCAGGTGTTTATCCCCGCCAGGC 876
Db 1055 TAGCTCCTTTATGCAGGCTCCCCGAGCAGCAGGAGATGGTGCAGGTGTTTATCCCCGCCAGGC 1114
Qy 877 AGTGGGGCGCCCATCATCGGCAAGAGAGGGCAGCAGCATCAAACAGCTCCTCCCGGTTGCCAG 936
Db 1115 AGTGGGGCGCCCATCATCGGCAAGAGAGGGCAGCAGCATCAAACAGCTCCTCCCGGTTGCCAG 1174
Qy 937 CGCCTCCATCAAGATTGCCACCACCCGAAACACCACTGACTCCAAAGTTCGTATGGTTATCAT 996
Db 1175 CGCCTCCATCAAGATTGCCACCACCCGAAACACCACTGACTCCAAAGTTCGTATGGTTATCAT 1234
Qy 997 CACTGGACCACGCGCCAGAGCCCAATTCAAGGCTCAGGAGAAGAATCTATGGCAAACTCAAGGA 1056
Db 1235 CACTGGACCACGCGCCAGAGCCCAATTCAAGGCTCAGGAGAAGAATCTATGGCAAACTCAAGGA 1294
Qy 1057 GGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGAGACACCCACATACGTGTGCCAGC 1116
Db 1295 GGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGAGACACCCACATACGTGTGCCAGC 1354
Qy 1117 ATCAGCAGCTGCGCCGGTCATTGGCAAAGGTGGAAAAACGTGAACCAGTTGCAGAATTT 1176
Db 1355 ATCAGCAGCTGCGCCGGTCATTGGCAAAGGTGGAAAAACGTGAACCAGTTCAGAATTT 1414
Qy 1177 GACCGGCAGCTGAGGTGGTAGTACCCAAGAACGACCAGGACGAACGACCAGGTCAT 1236
Db 1415 GACCGGCAGCTGAGGTGGTAGTACCCAAGAACGACCAGGACGAACGACCAGGTCAT 1474
Qy 1237 CGGTGAAAATCATCGGACATTTCTATCGCCAGTGCTCAACGAGATCCGAGACAT 1296
Db 1475 CGGTGAAAATCATCGGACATTTCTATGCCAGTGCTCAACGAGATCCGAGACAT 1534
Qy 1297 CCTGGGCCCAGGTTAAGCAGGGTCAGCAGAGGACATCAGAGGGACAGAGTAACCACGGAG 1356
Db 1535 CCTGGGCCCAGGTTAAGCAGGGTCAGCAGAGGACATCAGAGGGACAGAGTAACCACGGAG 1594
Qy 1357 GAAGTGACCAGCCCTCCCTGTCCCTTGAGTCCTTGAGTCCAGAAATCGGA 1416
Db 1595 GAAGTGACCAGCCCTCCCTGTCCCTTGAGTCCTTGAGTCCAGAAATCGGA 1654

Qy 1417 GTGTGCTCTCCCCGGCCAGGCCCTGAGAATGAGTGGGAATCCGGACACNTGGGCCGGGGCTG 1476
Db 1655 GTGTGCTCTCCCCGGCCAGGCCCTGAGAATGAGTGGGAATCCGGACACNTGGGCCGGGGCTG 1714
Qy 1477 TAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGAAAGATGTTCCAGTCA 1536
Db 1715 TAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGAAAGATGTTCCAGTCA 1774
Qy 1537 GCCCCAAACACCACCAATTGGCCAACACTGTNGCCCTCGGGTGTCAGAAATNT 1596
Db 1775 GCCCCAAACACCACCAATTGGCCAACACTGTNGCCCTCGGGTGTCAGAAATNT 1834
Qy 1597 AGCGGCAAGGCACTTTTAAACGTGGATTGTTTAAAGAAGTTCTCCAGGCCCCACCAAGAGG 1656
Db 1835 AGCGGCAAGGCACTTTTAAACGTGGATTGTTTAAAGAAGTTCTCCAGGCCCCACCAAGAGG 1894
Qy 1657 GTGGATCACACCTCAGTGGGAAGAAGAAAAATPAAAATTTCCTTCAGTTTTAAAA 1708
Db 1895 GTGGATCACACCTCAGTGGGAAGAAGAAAAATAAAAATTTCCTTCAGTTTTAAAA 1946

RESULT 4
US-09-899-651-7
; Sequence 7, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-7

  Query Match   97.9%; Score 1672; DB 4; Length 1946;
  Best Local Similarity 100.0%; Pred. No. 0;
  Matches 1672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   37 AGCCATCATGAAGTGAATGGCCACCAGTTGGAGACCATGCCCTGAAGGTCTCCTACAT 96
Db 275 AGCCATCATGAAGTGAATGGCCACCAGTTGGAGACCATGCCCTGAAGGTCTCCTACAT 334
Qy   97 CCCCGATGAGCAGATAGCACAGGGACCTGAGAATGGGCGCGCCAGCGGGTTTGGCTCTCG 156
Db 335 CCCCGATGAGCAGATAGCACAGGGACCTGAGAATGGGCGCGGGGCCAGCGGGTTTGGCTCTCG 394
Qy 157 GGGTCAGCCCCGCCAGGGGTCACCTGTGCAGCGGCCCAAGCCAGCAGCAAGT 216
Db 395 GGGTCAGCCCCGCCAGGGGTCACCTGTGCAGCGGCCCAAGCCAGCAGCAAGT 454
Qy 217 GGACATCCCCTTCGGCTCCTCCGTGCTGCCCACCCAGTGTGGGTGCCATTATTGGCAAGGA 276
Db 455 GGACATCCCCTTCGGCTCCTCCGTGCCACCCAGTATGTGGGTGCCATTATTGGCAAGGA 514
Qy 277 GGGGGCCACCATCCGCAACATCCAACATCACAAAACAGACCCAGTCCAAGATAGAGCTGCATAGGAA 336
Db 515 GGGGGCCACCATCCGCAACATCCAACATCACAAAACAGACCCAGTCCAAGATAGAGCTGCATAGGAA 574

Qy 1417 GTGTGCTCTCCCCGGCAGGCCTGAGAATGAGTGGGAATCCGGGACACNTGGGCGCGGGCTG 1476
         ||||||||||||||||||||||||||||||||| |||||||||||| ||||||||||||||
Db 1655 GTGTGCTCTCCCCGGCAGGCCTGAGAATGAGTGGGAATCCGGGACACNTGGGCGCGGGCTC 1714

Qy 1477 TAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGATCNTCA 1536
         |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db 1715 TAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGATCNTCA 1774

Qy 1537 GCCCCAAACACCCACCCAATTGGCCCAACACTGNTGCCCGGGTGTCAGAAATTNT 1596
         ||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db 1775 GCCCCAAACACCCACCCAATTGGCCCAACACTGNTGCCCGGGTGTCAGAAATTNT 1834

Qy 1597 AGGCAAGGCACTTTAAACGTGGATTGTTAAAGAAGCTCTCCAGGCCCCACCAAGAGG 1656
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1835 AGGCAAGGCACTTTAAACGTGGATTGTTAAAGAAGCTCTCCAGGCCCCACCAAGAGG 1894

Qy 1657 GTGGATCACACCTCAGTGGGAGAAGAAAATAAAATTCCTTCAGGTTTTAAAA 1708
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1895 GTGGATCACACCTCAGTGGGAGAAGAAAATAAAATTCCTTCAGGTTTTAAAA 1946

RESULT 5
US-09-261-855-1
; Sequence 1, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-261-855-1

Query Match   1.8%; Score 31; DB 3; Length 2224;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches   31; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   66 TGGAGAACCATGCCCTGAAGGTCTCCTACAT 96
         |||||||||||||||||||||||||||||||
Db 561 TGGAGAACCATGCCCTGAAGGTCTCCTACAT 591

RESULT 6
US-09-261-855-38/c
; Sequence 38, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: oligonucleotide primer
US-09-261-855-38

Query Match   1.4%; Score 24; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy 337 GGAGAACCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCACCCCTGAGGGCTGCTC 396
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 575 GGAGAACCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCACCCCTGAGGGCTGCTC 634

Qy 397 CTCCGCTTGTAAGATGATCTTGAGATTATGCATAAAGAGGCTAAGGACACCAAAAACGGC 456
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 635 CTCCGCTTGTAAGATGATCTTGAGATTATGCATAAAGAGGCTAAGGACACCAAAAACGGC 694

Qy 457 TGACGAGGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAA 516
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 695 TGACGAGGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAA 754

Qy 517 GGAAGGACGGAACCTGAGAAGGTAGACAAGATACCGAGACAAAAATCACCATCTCCTC 576
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 755 GGAAGGACGGAACCTGAGAAGGTAGAGCAAGATACCGAGACACAAAAATCACCATCTCCTC 814

Qy 577 GTTGCAAGACCTTACCCTTTACACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA 636
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 815 GTTGCAAGACCTTACCCTTTACACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA 874

Qy 637 GAATTGTTGCAGGGCCGAGCCAGGAGAAATAATGAAGAAGTTCGGGAGGCCTATGAGAATGA 696
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 875 GAATTGTTGCAGGGCCGAGCCAGGAGAAATAATGAAGAAGTTCGGGAGGCCTATGAGAATGA 934

Qy 697 TGTGGCTGCCATCATCGGCAAGAGGGGCAGCACCATCAAACAGCTCTCCCGGTTGCCAG 756
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 935 TGTGGCTGCCATCATCGGCAAGAGGGGCAGCACCATCAAACAGCTCTCCCGGTTGCCAG 994

Qy 757 CCCAGCTTCATCCAGCGCAGTCCCGCGCCTCCCAGCAGCGTTACTGGGGCTGCTCCCTA 816
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 995 CCCAGCTTCATCCAGCGCAGTCCCGCGCCTCCCAGCAGCGTTACTGGGGCTGCTCCCTA 1054

Qy 817 TAGCTCCTTATGCAGGGTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCCAGGC 876
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1055 TAGCTCCTTATGCAGGGTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCCAGGC 1114

Qy 877 AGTGGGCCGCCATCATCGGCAAGAGGGGCAGCACCATCAAACAGCTCTCCCGGTTTGCCAG 936
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1115 AGTGGGCCGCCATCATCGGCAAGAGGGGCAGCACCATCAAACAGCTCTCCCGGTTTGCCAG 1174

Qy 937 CGGCTCCATCAAGATTCGACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCAT 996
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1175 CGGCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCAT 1234

Qy 997 CACTGGACACGGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATGCAAACTCAAGGA 1056
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1235 CACTGGACACGGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATGCAAACTCAAGGA 1294

Qy 1057 GGAGAACTTCTTTGGTCCCAAGAGGAGGAGCTGAAGCTGAAGTGGAGCTGAAGCTGTGCCAGC 1116
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1295 GGAGAACTTCTTTGGTCCCAAGAGGAGGAGCTGAAGCTGAAGTGGAGCTGAAGCTGTGCCAGC 1354

Qy 1117 ATCAGCAGCTGGCGGGTCATTGGCAAGGTCCAAGTGGAAAAACGGTGAACGAGTTGCAGAATT 1176
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1355 ATCAGCAGCTGGCGGGTCATTGGCAAGGTGCAAAAACGGTGAACGAGTTGCAGAATT 1414

Qy 1177 GACGGCGAGCTGAGGTGGTAGTAGTGTCAAGAGAACGACCAGGTCAT 1236
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 1415 GACGGCGAGCTGAGGTGGTAGTAGTGTCAAGAGAACGACCAGGTCAT 1474

Qy 1237 CGTGAAAATCATCGGACAATTCTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACAT 1296
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1475 CGTGAAAATCATCGGACAATTCTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACAT 1534

Qy 1297 CCTGGCCCAGGTTAAGCAGCAGCATCAGAGGGACAGAGTAACCAGGCCCAGGCACGGAG 1356
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1535 CCTGGCCCAGGTTAAGCAGCAGCATCAGAGGGACAGAGTAACCAGGCCCAGGCACGGAG 1594

Qy 1357 GAAGTGACCAGCAGCCCCGTCCCTTGAGTCCAGGACAACAACGGCAGAAATCGAGA 1416
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1595 GAAGTGACCAGCAGCCCCGTCCCTTGAGTCCAGGACAACAACGGCAGAAATCGAGA 1654

```
Qy      508  CATTGGCAAGGAGGAGGACGGAACCT  531
             |||||||||||||||||||||||||||
Db       24  CATTGGCAAGGAGGAGGACGGAACCT  1

RESULT 7
US-09-261-855-36
; Sequence 36, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
;   APPLICANT: Ross, Jeffrey
;   TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
;   TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
;   FILE REFERENCE: 960296.95131
;   CURRENT APPLICATION NUMBER: US/09/261,855A
;   CURRENT FILING DATE: 1999-03-03
;   NUMBER OF SEQ ID NOS: 46
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide primer
US-09-261-855-36

Query Match             1.3%;  Score 23;  DB 3;  Length 23;
Best Local Similarity  100.0%;  Pred. No. 0.1;
Matches   23;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1178  ACGGCAGCTGAGGTGGTAGTACC  1200
             |||||||||||||||||||||||
Db        1  ACGGCAGCTGAGGTGGTAGTACC  23

RESULT 8
US-09-643-597-347
; Sequence 347, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Li, Samuel X.
;   APPLICANT: Wang, Aijun
;   APPLICANT: Skeiky, Yasir A.W.
;   APPLICANT: Henderson, Robert A.
;   APPLICANT: McNeill, Patricia A.
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C11
;   CURRENT APPLICATION NUMBER: US/09/643,597
;   CURRENT FILING DATE: 2000-08-21
;   NUMBER OF SEQ ID NOS: 369
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 347
;   LENGTH: 1740
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-643-597-347

Query Match             1.2%;  Score 20;  DB 4;  Length 1740;
Best Local Similarity  100.0%;  Pred. No. 4;
Matches   20;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1019  TTCAAGGCTCAGGGAAGAAT  1038
             ||||||||||||||||||||
Db     1390  TTCAAGGCTCAGGGAAGAAT  1409
```

```
RESULT 9
US-09-542-615A-347
; Sequence 347, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy A.
;   APPLICANT: Fanger, Gary R.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C8
;   CURRENT APPLICATION NUMBER: US/09/542,615A
;   CURRENT FILING DATE: 2000-04-14
;   NUMBER OF SEQ ID NOS: 350
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 347
;   LENGTH: 1740
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-542-615A-347

Query Match             1.2%;  Score 20;  DB 4;  Length 1740;
Best Local Similarity  100.0%;  Pred. No. 4;
Matches   20;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1019  TTCAAGGCTCAGGGAAGAAT  1038
             ||||||||||||||||||||
Db     1390  TTCAAGGCTCAGGGAAGAAT  1409

RESULT 10
US-09-606-421B-347
; Sequence 347, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Li, Samuel X.
;   APPLICANT: Wang, Aijun
;   APPLICANT: Skeiky, Yasir A.W.
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C9
;   CURRENT APPLICATION NUMBER: US/09/606,421B
;   CURRENT FILING DATE: 2000-06-28
;   NUMBER OF SEQ ID NOS: 358
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 347
;   LENGTH: 1740
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-606-421B-347

Query Match             1.2%;  Score 20;  DB 4;  Length 1740;
Best Local Similarity  100.0%;  Pred. No. 4;
Matches   20;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1019  TTCAAGGCTCAGGGAAGAAT  1038
             ||||||||||||||||||||
Db     1390  TTCAAGGCTCAGGGAAGAAT  1409

RESULT 11
US-09-061-709-4
; Sequence 4, Application US/09061709B
```

; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-4

Query Match   1.2%;   Score 20;   DB 3;   Length 4159;
Best Local Similarity   100.0%;   Pred. No. 4.2;
Matches   20;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1019 TTCAAGGCTCAGGGAAGAAT 1038
            ||||||||||||||||||||
Db   1640 TTCAAGGCTCAGGGAAGAAT 1659

RESULT 12
US-09-899-651-4
; Sequence 4, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-4

Query Match   1.2%;   Score 20;   DB 4;   Length 4159;
Best Local Similarity   100.0%;   Pred. No. 4.2;
Matches   20;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1019 TTCAAGGCTCAGGGAAGAAT 1038
            ||||||||||||||||||||
Db   1640 TTCAAGGCTCAGGGAAGAAT 1659

RESULT 13
US-09-643-597-175
; Sequence 175, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3502)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3549)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3940)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3968)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3974)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4036)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4056)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4062)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4080)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A,T,C or G
US-09-643-597-175

Query Match   1.2%;   Score 20;   DB 4;   Length 4181;
Best Local Similarity   100.0%;   Pred. No. 4.2;
Matches   20;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1019 TTCAAGGCTCAGGGAAGAAT 1038
           ||||||||||||||||||||
Db   1640 TTCAAGGCTCAGGGAAGAAT 1659

RESULT 14
US-09-480-884A-175
; Sequence 175, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Hosken, Nancy A.
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Fanger, Gary R.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C6
;   CURRENT APPLICATION NUMBER: US/09/480,884A
;   CURRENT FILING DATE: 2001-08-27
;   NUMBER OF SEQ ID NOS: 330
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 175
;   LENGTH: 4181
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(4181)
;   OTHER INFORMATION: n=A,T,C or G
US-09-480-884A-175

Query Match   1.2%;   Score 20;   DB 4;   Length 4181;
Best Local Similarity   100.0%;   Pred. No. 4.2;
Matches   20;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1019 TTCAAGGCTCAGGGAAGAAT 1038
           ||||||||||||||||||||
Db   1640 TTCAAGGCTCAGGGAAGAAT 1659

RESULT 15
US-09-542-615A-175
; Sequence 175, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy A.
;   APPLICANT: Fanger, Gary R.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C8
;   CURRENT APPLICATION NUMBER: US/09/542,615A
;   CURRENT FILING DATE: 2000-04-14
;   NUMBER OF SEQ ID NOS: 350
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 175
;   LENGTH: 4181
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (3347)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3502)

;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3506)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3520)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3538)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3549)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3646)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3940)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3968)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3974)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4036)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4056)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4062)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4080)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4088)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4115)
;   OTHER INFORMATION: n=A,T,C or G
US-09-542-615A-175

Query Match   1.2%;   Score 20;   DB 4;   Length 4181;
Best Local Similarity   100.0%;   Pred. No. 4.2;
Matches   20;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   1019 TTCAAGGCTCAGGGAAGAAT 1038
           ||||||||||||||||||||
Db   1640 TTCAAGGCTCAGGGAAGAAT 1659

Search completed: July 14, 2004, 03:05:05
Job time : 141 secs

```
             GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:      July 13, 2004, 12:14:48 ; Search time 19.5 Seconds
                                      (without alignments)
                                      9121.623 Million cell updates/sec

Title:       US-09-270-437D-5
Perfect score:  560
Sequence:    1 aggacggctgccgcaccgcc.........atttccttcaggttttaaaa 1708

Scoring table:  OLIGO
             Xgapop 60.0 , Xgapext 60.0
             Ygapop 60.0 , Ygapext 60.0
             Fgapop  6.0 , Fgapext  7.0
             Delop   6.0 , Delext   7.0

Searched:    141681 seqs, 52070155 residues

Word size:   1

Total number of hits satisfying chosen parameters:    283186

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_13072004_121955_9970/app_query.fasta_1.1863
-DB=SwissProt_42 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN_1_1_24 @runat_13072004_121955_9970 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :   SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	8	1.4	166	1	ING_CAMBA	Q865w6 camelus bac
2	8	1.4	166	1	ING_LAMGL	Q865x1 lama glama
3	8	1.4	224	1	QCRB_BACTC	Q45658 bacillus th
4	8	1.4	227	1	YVEL_BACSU	P71051 bacillus su
5	8	1.4	259	1	YCBC_ECOLI	P36565 escherichia
6	8	1.4	267	1	RS2_DROME	P31009 drosophila
7	8	1.4	371	1	GP8_BRARE	Q91919 brachydanio
8	8	1.4	473	1	PHDK_NOCSK	O24723 nocardioide
9	8	1.4	512	1	DFN5_MOUSE	Q9z2d3 mus musculu
10	8	1.4	600	1	Y018_MYCGE	P47264 mycoplasma
11	8	1.4	646	1	SG1_BOVIN	P23389 bos taurus
12	8	1.4	739	1	BACI_MOUSE	P97302 mus musculu
13	8	1.4	815	1	GYRB_MYXXA	O33367 myxococcus
14	8	1.4	1030	1	Y018_MYCPN	P75093 mycoplasma
15	8	1.4	1473	1	TOP2_ARATH	P30182 arabidopsis
16	8	1.4	1648	1	YJ9H_YEAST	P47171 saccharomyc
17	8	1.4	2567	1	M18B_HUMAN	Q8iug5 homo sapien
18	8	1.4	3176	1	CA36_HUMAN	P12111 homo sapien
19	8	1.4	4451	1	GRSB_BACBR	P14688 b gramicidi
20	8	1.4	505	1	RYR1_PIG	P16960 sus scrofa
21	8	1.4	5037	1	RYR1_RABIT	P11716 oryctolagus
22	8	1.4	5038	1	RYR1_HUMAN	P21817 homo sapien
23	7	1.4	55	1	RL32_VIBCH	Q9kqh3 vibrio chol
24	7	1.2	85	1	RM33_YEAST	P20084 saccharomyc
25	7	1.3	85	1	FIS_ERWCA	052540 erwinia car
26	7	1.3	98	1	FIS_YERPE	Q8zax8 yersinia pe
27	7	1.3	98	1	CBGR_CLOAB	P24132 clostridium
28	7	1.3	101	1	LV2J_HUMAN	P01713 homo sapien
29	7	1.3	110	1	GVK1_HALN1	P24375 halobacteri
30	7	1.3	113	1	RLA2_BRUMA	P90703 brugia mala
31	7	1.3	114	1	VA31_VACCC	P21096 vaccinia vi
32	7	1.3	124	1	VA31_VACCV	P24760 vaccinia vi
33	7	1.3	136	1	C17_HUMAN	Q9nrrl homo sapien
34	7	1.2	136	1	FLIS_VIBCH	Q9kq65 vibrio chol
35	7	1.3	140	1	VA31_VARV	P33848 variola vir
36	7	1.3	149	1	GLPA_PANTR	Q28913 pan troglod
37	7	1.2	150	1	ME15_SCHPO	Q96ws6 schizosacch
38	7	1.2	153	1	HS11_SOYBN	P02519 glycine max
39	7	1.3	164	1	ING_CHICK	P49708 gallus gall
40	7	1.3	164	1	ING_COTJA	057571 coturnix co
41	7	1.3	164	1	ING_MELGA	057603 meleagris g
42	7	1.3	164	1	ING_NUMME	073915 numida mele
43	7	1.3	164	1	ING_PHACO	057608 phasianus c
44	7	1.3	165	1	GRPI_ORYSA	P25074 oryza sativ
45	7	1.3	173	1	HA34_BRELC	Q99074 bremia lact

ALIGNMENTS

```
RESULT 1
ING_CAMBA
ID   ING_CAMBA     STANDARD;      PRT;    166 AA.
AC   Q865W6;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Interferon gamma precursor (IFN-gamma).
GN   IFNG.
OS   Camelus bactrianus (Bactrian camel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX   NCBI_TaxID=9837;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Raadan O., Lee S., Yoshida R., Chang K., Ohashi K., Sugimoto C.,
RA   Onuma M.;
RT   "Cloning and sequence analysis of cytokine cDNAs of llama and camel.";
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Produced by lymphocytes activated by specific antigens
CC       or mitogens. IFN-gamma, in addition to having antiviral activity,
CC       has important immunoregulatory functions. It is a potent activator
CC       of macrophages, it has antiproliferative effects on transformed
CC       cells and it can potentiate the antiviral and antitumor effects of
CC       the type I interferons (By similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Released primarily from activated T
CC       lymphocytes.
CC   -!- SIMILARITY: Belongs to the type II (or gamma) interferon family.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; AB107657; BAC75394.1; -.
DR   InterPro; IPR002069; IFN-gamma.
DR   Pfam; PF00714; IFN-gamma; 1.
```

DR   ProDom; PD002435; IFN-gamma; 1.
KW   Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.
FT   SIGNAL   1   20   BY SIMILARITY.
FT   CHAIN   21   166   INTERFERON GAMMA.
FT   CARBOHYD   39   39   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   106   106   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   166 AA; 19402 MW; 4547BC4FC693655 CRC64;

Alignment Scores:
Pred. No.:   36.5   Length: 166
Score: 8.00   Matches: 8
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismacches: 0
Query Match: 1.44%   Indels: 0
DB: Gaps: 0

US-09-270-437D-5 (1-1708) x ING_CAMBA (1-166)

Qy 298 GATGTTGCGGATGGTGGCCCCCTC 275
        ||||||||||||||||||||||||
Db   44 AspValAlaAspGlyGlyProLeu 51

RESULT 2
ING_LAMGL
ID   ING_LAMGL   STANDARD;   PRT;   166 AA.
AC   Q865X1;
DT   15-MAR-2004 (Rel. 43, Created)
DT   15-MAR-2004 (Rel. 43, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Interferon gamma precursor (IFN-gamma).
GN   IFNG.
OS   Lama glama (llama).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
OX   NCBI_TaxID=9844;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Raadan O., Lee S., Yoshida R., Chang K., Ohashi K., Sugimoto C.,
RA   Onuma M.;
RT   "Cloning and sequence analysis of cytokine cDNAs of llama and camel.";
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Produced by lymphocytes activated by specific antigens
CC   or mitogens. IFN-gamma, in addition to having antiviral activity,
CC   has important immunoregulatory functions. It is a potent activator
CC   of macrophages, it has antiproliferative effects on transformed
CC   cells and it can potentiate the antiviral and antitumor effects of
CC   the type I interferons (By similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Released primarily from activated T
CC   lymphocytes.
CC   -!- SIMILARITY: Belongs to the type II (or gamma) interferon family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AB107652; BAC75389.1; -.
DR   InterPro; IPR002069; IFN-gamma.
DR   Pfam; PF00714; IFN-gamma; 1.
DR   ProDom; PD002435; IFN-gamma; 1.
KW   Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.
FT   SIGNAL   1   20   BY SIMILARITY.
FT   CHAIN   21   166   INTERFERON GAMMA.
FT   CARBOHYD   39   39   N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD   106   106   N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   166 AA; 19475 MW; 23937B814759328F CRC64;

Alignment Scores:

Pred. No.:   36.5   Length: 166
Score: 8.00   Matches: 8
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 1.44%   Indels: 0
DB: Gaps: 0

US-09-270-437D-5 (1-1708) x ING_LAMGL (1-166)

Qy 298 GATGTTGCGGATGGTGGCCCCCTC 275
        ||||||||||||||||||||||||
Db   44 AspValAlaAspGlyGlyProLeu 51

RESULT 3
QCRB_BACTC
ID   QCRB_BACTC   STANDARD;   PRT;   224 AA.
AC   Q45658;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Menaquinol-cytochrome c reductase cytochrome B subunit.
GN   QCRB.
OS   Bacillus thermodenitrificans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX   NCBI_TaxID=33940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K1041;
RX   MEDLINE=96218169; PubMed=8647852;
RA   Sone N., Tsuchiya N., Inoue M., Noguchi S.;
RT   "Bacillus stearothermophilus qcr operon encoding Rieske FeS protein,
RT   cytochrome b6, and a novel-type cytochrome c1 of quinol-cytochrome c
RT   reductase.";
RL   J. Biol. Chem. 271:12457-12462(1996).
CC   -!- FUNCTION: Component of the menaquinol-cytochrome c reductase
CC   complex.
CC   -!- COFACTOR: Two heme groups which are not covalently bound to the
CC   protein (By similarity).
CC   -!- SUBUNIT: THE MAIN SUBUNITS OF THE MENAQUINONE:CYTOCHROME C COMPLEX
CC   ARE: CYTOCHROME B, THE RIESKE PROTEIN AND A 22/29 kDa CYTOCHROME
CC   B/C SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY. CORRESPONDS TO THE
CC   AMINO END OF MITOCHONDRIAL CYTOCHROME B.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; D83789; BAA12117.1; -.
DR   InterPro; IPR005797; Cytb_b6_N.
DR   Pfam; PF00033; cytochrome_b_N; 1.
DR   PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
KW   Electron transport; Heme; Transmembrane.
FT   TRANSMEM   37   57   POTENTIAL.
FT   TRANSMEM   96   116   POTENTIAL.
FT   TRANSMEM   126   146   POTENTIAL.
FT   TRANSMEM   195   215   POTENTIAL.
FT   METAL   94   94   IRON (HEME 1 AXIAL LIGAND)
FT   (BY SIMILARITY).
FT   METAL   108   108   IRON (HEME 2 AXIAL LIGAND)
FT   (BY SIMILARITY).
FT   METAL   196   196   IRON (HEME 1 AXIAL LIGAND)
FT   (BY SIMILARITY).
FT   METAL   211   211   IRON (HEME 2 AXIAL LIGAND)
FT   (BY SIMILARITY).
SQ   SEQUENCE   224 AA; 25411 MW; AD05E81F4E7F0518 CRC64;

Alignment Scores:

```
Pred. No.:              35.1          Length:          224
Score:                  8.00          Matches:         8
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            1.43%         Indels:          0
DB:                                   Gaps:            0

US-09-270-437D-5 (1-1708) x QCRB_BACTC (1-224)

Qy    786  CTCCCAGCAGCGTTACTGGGGCTG 809
           ||||||||||||||||||||||||
Db    200  LeuProAlaAlaLeuLeuGlyLeu 207

RESULT 4
YVEL_BACSU
ID   YVEL_BACSU      STANDARD;      PRT;      227 AA.
AC   P71051; O08170;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Putative tyrosine-protein kinase yveL (EC 2.7.1.112).
GN   YVEL OR BSU34360.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=97124193; PubMed=8969506;
RA   Fabret C., Quentin Y., Chapal N., Guiseppi A., Haiech J., Denizot F.;
RT   "Integrated mapping and sequencing of a 115 kb DNA fragment from
RT   Bacillus subtilis: sequence analysis of a 21 kb segment containing
RT   the sigL locus.";
RL   Microbiology 142:3089-3096(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Denizot F.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Duesterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC       tyrosine phosphate.
CC   -!- SIMILARITY: Belongs to the cpsD/capB family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z71928; CAA96490.1; -.
DR   EMBL; Z94043; CAB08024.1; -.
DR   EMBL; Z99121; CAB15441.1; -.
DR   PIR; A70036; A70036.
DR   Subtilist; BG11860; yveL.
DR   InterPro; IPR005702; EPS_synthesis.
DR   TIGRFAMs; TIGR01007; eps_fam; 1.
KW   Hypothetical protein; Transferase; Tyrosine-protein kinase;
KW   Complete proteome.
SQ   SEQUENCE 227 AA;  24674 MW;  81CBB9D75278FFA9 CRC64;

Alignment Scores:
Pred. No.:              35.1          Length:          227
Score:                  8.00          Matches:         8
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            1.44%         Indels:          1
DB:                                   Gaps:            0

US-09-270-437D-5 (1-1708) x YVEL_BACSU (1-227)

Qy   1367  CTGGTCACTTCCTCCGTGCCTGG 1344
           ||||||||||||||||||||||||
Db     48  LeuValThrSerSerValProGly 55

RESULT 5
YCBC_ECOLI
ID   YCBC_ECOLI      STANDARD;      PRT;      259 AA.
AC   P36565; P75846;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein ycbC.
GN   YCBC OR B0920 OR Z1267 OR ECS1003.
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
```

RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
RN   [5]
RP   SEQUENCE OF 1-170 FROM N.A.
RC   STRAIN=K12 / W3110;
RX   MEDLINE=94232180; PubMed=7513784;
RA   Feng J., Yamanaka K., Niki H., Ogura T., Hiraga S.;
RT   "New killing system controlled by two genes located immediately
RT   upstream of the mukB gene in Escherichia coli.";
RL   Mol. Gen. Genet. 243:136-147(1994).
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AE000194; AAC74006.1; -.
DR   EMBL; D90730; BAA35666.1; -.
DR   EMBL; AE005281; AAG55405.1; -.
DR   EMBL; AP002553; BAB34426.1; -.
DR   EMBL; D26440; -; NOT_ANNOTATED_CDS.
DR   PIR; A85618; A85618.
DR   PIR; C90754; C90754.
DR   PIR; G64831; G64831.
DR   EcoGene; EG12166; ycbC.
DR   InterPro; IPR003848; DUF218.
DR   Pfam; PF02698; DUF218; 1.
KW   Hypothetical protein; Complete proteome.
FT   CONFLICT   89   89   Y -> N (IN REF. 5).
FT   CONFLICT   150   171   GVPREQIITLDLIPKDTEEEAAA -> ACRASKLSPWICQKI
FT   PKKKLQ (IN REF. 5).
SQ   SEQUENCE   259 AA;   28666 MW;   7EAE14C696DAA0C9 CRC64;

Alignment Scores:
Pred. No.:
Score: 34.5
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.44%
DB:

Length: 259
Matches: 8.00
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-270-437D-5 (1-1708) x YCBC_ECOLI (1-259)

Qy   479   GGATCTTCAGGGGAACCTCGTCAG   456
          ||||||||||||||||||||||||
Db   251   GlySerSerGlyGluProArgGln   258

RESULT 6
RS2_DROME
ID   RS2_DROME   STANDARD;   PRT;   267 AA.
AC   P31009; Q9VL74;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   40S ribosomal protein S2 (Strings of pearls protein).
GN   SOP OR RPS2 OR CG5920.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95073591; PubMed=7982558;
RA   Cramton S.E., Laski F.A.;
RT   "String of pearls encodes Drosophila ribosomal protein S2, has
RT   Minute-like characteristics, and is required during oogenesis.";
RL   Genetics 137:1039-1048(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   SEQUENCE OF 1-241 FROM N.A.
RC   STRAIN=Canton-S;
RX   MEDLINE=91811212; PubMed=8441641;
RA   Barrio R., del Arco A., Cabrera H.L., Arribas C.;
RT   "Cloning and analysis of the S2 ribosomal protein cDNA from
RT   Drosophila.";
RL   Nucleic Acids Res. 21:351-351(1993).
CC   -!- SIMILARITY: Belongs to the S5P family of ribosomal proteins.
CC   -!- SIMILARITY: Contains 1 S5 DRBM domain.

```
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X69120; CAA48872.1; -.
DR   EMBL; AE003626; AAF52822.1; -.
DR   EMBL; U01334; AAC34198.1; -.
DR   EMBL; U01335; AAA87053.1; -.
DR   PIR; S30395; S30395.
DR   HSSP; P02357; 1PKP.
DR   FlyBase; FBgn0004867; sop.
DR   InterPro; IPR000851; Ribosomal_S5.
DR   InterPro; IPR003324; Ribosomal_S5_C.
DR   InterPro; IPR005711; Ribosomal_S5_e/a.
DR   Pfam; PF00333; Ribosomal_S5; 1.
DR   Pfam; PF03719; Ribosomal_S5_C; 1.
DR   TIGRFAMs; TIGR01020; rpsE_arch; 1.
DR   PROSITE; PS00585; RIBOSOMAL_S5; 1.
DR   PROSITE; PS50881; S5_DSRBD; 1.
KW   Ribosomal protein.
FT   DOMAIN       85    148       S5 DRBM.
FT   CONFLICT     19     20       GG -> PP (IN REF. 3).
FT   CONFLICT    194    194       K -> R (IN REF. 3).
SQ   SEQUENCE   267 AA;  28899 MW;  ADA22CD28F100743 CRC64;

Alignment Scores:
Pred. No.:                     34.4         Length:             267
Score:                         8.00         Matches:              8
Percent Similarity:         100.00%         Conservative:         0
Best Local Similarity:      100.00%         Mismatches:           0
Query Match:                  1.43%         Indels:               0
DB:                                         Gaps:                 0

US-09-270-437D-5 (1-1708) x RS2_DROME (1-267)

Qy      137 CGAGGGGGCTTTGGCTCTCGGGGT 160
            ||||||||||||||||||||||||
Db       12 ArgGlyGlyPheGlySerArgGly 19

RESULT 7
GP85_BRARE
ID   GP85_BRARE      STANDARD;      PRT;    371 AA.
AC   Q91919;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Super conserved receptor expressed in brain 2.
GN   SREB2.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20294882; PubMed=10833454;
RA   Matsumoto M., Saito T., Takasaki J., Kamohara M., Sugimoto T.,
RA   Kobayashi M., Tadokoro M., Matsumoto S., Ohishi T., Furuichi K.;
RT   "An evolutionarily conserved G-protein coupled receptor family, SREB,
RT   expressed in the central nervous system.";
RL   Biochem. Biophys. Res. Commun. 272:576-582(2000).
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AB040805; BAA96651.1; -.
DR   ZFIN; ZDB-GENE-000710-2; sreb2.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family.
FT   DOMAIN        1     26       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     27     47       1 (POTENTIAL).
FT   DOMAIN       48     58       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     59     79       2 (POTENTIAL).
FT   DOMAIN       80     96       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     97    117       3 (POTENTIAL).
FT   DOMAIN      118    138       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    139    159       4 (POTENTIAL).
FT   DOMAIN      160    189       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    190    210       5 (POTENTIAL).
FT   DOMAIN      211    287       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    288    308       6 (POTENTIAL).
FT   DOMAIN      309    321       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    322    342       7 (POTENTIAL).
FT   DOMAIN      343    371       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     95    173       BY SIMILARITY.
FT   CARBOHYD      3      3       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     84     84       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    183    183       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   371 AA;  41954 MW;  F6F6175ED3A348C2 CRC64;

Alignment Scores:
Pred. No.:                       33         Length:             371
Score:                         8.00         Matches:              8
Percent Similarity:         100.00%         Conservative:         0
Best Local Similarity:      100.00%         Mismatches:           0
Query Match:                  1.44%         Indels:               0
DB:                                         Gaps:                 0

US-09-270-437D-5 (1-1708) x GP85_BRARE (1-371)

Qy      802 AGTAACGCTGCTGGGAGGCGGCGG 779
            ||||||||||||||||||||||||
Db      264 SerAsnAlaAlaGlyArgArgArg 271

RESULT 8
PHDK_NOCSK
ID   PHDK_NOCSK      STANDARD;      PRT;    473 AA.
AC   O24723;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Probable 1-hydroxy-2-naphthoate transporter.
OS   Nocardioides sp. (strain KP7).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Propionibacterineae; Nocardioidaceae; Nocardioides.
OX   NCBI_TaxID=35761;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97474276; PubMed=9335300;
RA   Iwabuchi T., Harayama S.;
RT   "Biochemical and genetic characterization of 2-carboxybenzaldehyde
RT   dehydrogenase, an enzyme involved in phenanthrene degradation by
RT   Nocardioides sp. strain KP7.";
RL   J. Bacteriol. 179:6488-6494(1997).
CC   -!- FUNCTION: PROBABLY INVOLVED IN THE TRANSPORT OF 1-HYDROXY-2-
CC       NAPHTHOATE.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (Potential).
```

```
CC    -!- SIMILARITY: Belongs to the sugar transporter family.
CC    ---------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------
DR    EMBL; AB000735; BAA23264.1; -.
DR    InterPro; IPR007114; MFS.
DR    InterPro; IPR005828; Sub_transporter.
DR    InterPro; IPR005829; Sug_transporter.
DR    Pfam; PF00083; sugar_tr; 1.
DR    PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR    PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
KW    Transport; Transmembrane; Inner membrane.
FT    DOMAIN         1     25       CYTOPLASMIC (POTENTIAL).
FT    TRANSMEM      26     46       1 (POTENTIAL).
FT    DOMAIN        47     59       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM      60     80       2 (POTENTIAL).
FT    DOMAIN        81     91       CYTOPLASMIC (POTENTIAL).
FT    TRANSMEM      92    112       3 (POTENTIAL).
FT    DOMAIN       113    121       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM     122    142       4 (POTENTIAL).
FT    DOMAIN       143    153       CYTOPLASMIC (POTENTIAL).
FT    TRANSMEM     154    174       5 (POTENTIAL).
FT    DOMAIN       175    180       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM     181    201       6 (POTENTIAL).
FT    DOMAIN       202    263       CYTOPLASMIC (POTENTIAL).
FT    TRANSMEM     264    284       7 (POTENTIAL).
FT    DOMAIN       285    304       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM     305    325       8 (POTENTIAL).
FT    DOMAIN       326    330       CYTOPLASMIC (POTENTIAL).
FT    TRANSMEM     331    351       9 (POTENTIAL).
FT    DOMAIN       352    354       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM     355    375       10 (POTENTIAL).
FT    DOMAIN       376    398       CYTOPLASMIC (POTENTIAL).
FT    TRANSMEM     399    418       11 (POTENTIAL).
FT    DOMAIN       419    421       EXTRACELLULAR (POTENTIAL).
FT    TRANSMEM     422    444       12 (POTENTIAL).
FT    DOMAIN       445    473       CYTOPLASMIC (POTENTIAL).
SQ    SEQUENCE   473 AA;  49109 MW;  D6D765D376260D8A CRC64;

Alignment Scores:
Pred. No.:                 32          Length:            473
Score:                      8.00       Matches:             8
Percent Similarity:       100.00%      Conservative:        0
Best Local Similarity:    100.00%      Mismatches:          0
Query Match:                1.44%      Indels:              0
DB:                                    Gaps:                0

US-09-270-437D-5 (1-1708) x PHDK_NOCSK (1-473)

Qy     1217 CAGGGGTTCTGGTCTCTTGGTACTA 1194
               |||||||||||||||||||||||||
Db      277 GlnGlySerGlyLeuLeuValLeu 284

RESULT 9
DFN5_MOUSE
ID    DFN5_MOUSE    STANDARD;    PRT;    512 AA.
AC    Q9Z2D3;
DT    16-OCT-2001 (Rel. 40, Created)
DT    16-OCT-2001 (Rel. 40, Last sequence update)
DT    16-OCT-2001 (Rel. 40, Last annotation update)
DE    Nonsyndromic hearing impairment protein 5 homolog.
GN    DFNA5 OR DFNA5H.
OS    Mus musculus (Mouse).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX    NCBI_TaxID=10090;
```

```
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Cochlea;
RX    MEDLINE=98442658; PubMed=9771715;
RA    Van Laer L., Huizing E.H., Verstreken M., van Zuijlen D.,
RA    Wauters J.G., Bossuyt P.J., Van de Heyning P., McGuirt W.T.,
RA    Smith R.J.H., Willems P.J., Legan P.K., Richardson G.P., Van Camp G.;
RT    "Nonsyndromic hearing impairment is associated with a mutation in
RT    DFNA5.";
RL    Nat. Genet. 20:194-197(1998).
CC    -!- SIMILARITY: BELONGS TO THE DFN5 FAMILY.
CC    ---------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------
DR    EMBL; AF073309; AAC69325.1; -.
DR    MGD; MGI:1889850; Dfna5h.
DR    InterPro; IPR007677; DFNA5.
DR    InterPro; IPR007681; Mob1.
DR    Pfam; PF04598; DFNA5; 1.
SQ    SEQUENCE   512 AA;  56630 MW;  13AFB8627773C4A5 CRC64;

Alignment Scores:
Pred. No.:                 31.7        Length:            512
Score:                      8.00       Matches:             8
Percent Similarity:       100.00%      Conservative:        0
Best Local Similarity:    100.00%      Mismatches:          0
Query Match:                1.44%      Indels:              0
DB:                                    Gaps:                0

US-09-270-437D-5 (1-1708) x DFN5_MOUSE (1-512)

Qy      616 GATGGTCCCTCAGGGTTGTAAAG 593
               ||||||||||||||||||||||
Db      279 AspGlyProLeuArgValValLys 286

RESULT 10
Y018_MYCGE
ID    Y018_MYCGE    STANDARD;    PRT;    600 AA.
AC    P47264; Q49209; Q49302;
DT    01-FEB-1996 (Rel. 33, Created)
DT    15-JUL-1999 (Rel. 38, Last sequence update)
DT    16-OCT-2001 (Rel. 40, Last annotation update)
DE    Hypothetical helicase MG018.
GN    MG018.
OS    Mycoplasma genitalium.
OC    Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX    NCBI_TaxID=2097;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=ATCC 33530 / G-37;
RX    MEDLINE=96026346; PubMed=7569993;
RA    Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA    Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA    Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA    Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA    Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA    Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT    "The minimal gene complement of Mycoplasma genitalium.";
RL    Science 270:397-403(1995).
RN    [2]
RP    REVISIONS.
RA    Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA    Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA    Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA    Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA    Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA    Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,
```

```
RL    Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN    [3]
RP    SEQUENCE OF 209-309 AND 371-471 FROM N.A.
RC    STRAIN=ATCC 33530 / G-37;
RX    MEDLINE=94075230; PubMed=8253680;
RA    Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT    "A survey of the Mycoplasma genitalium genome by using random
RT    sequencing.";
RL    J. Bacteriol. 175:7918-7930(1993).
CC    -!- MISCELLANEOUS: IN M.PNEUMONIAE, A SINGLE ORF SPANS M.GENITALIUM
CC        MG016, MG017 AND MG018.
CC    -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
CC    -------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    -------------------------------------------------------------------
DR    EMBL; U39681; AAC71234.1; -.
DR    EMBL; U02179; AAD12465.1; -.
DR    EMBL; U01757; AAD10571.1; -.
DR    PIR; T09676; T09676.
DR    TIGR; MG018; -.
DR    InterPro; IPR001410; DEAD.
DR    InterPro; IPR001650; Helicase_C.
DR    InterPro; IPR000330; SNF2_N.
DR    Pfam; PF00271; helicase_C; 1.
DR    Pfam; PF00176; SNF2_N; 1.
DR    SMART; SM00487; DEXDc; 1.
DR    SMART; SM00490; HELICc; 1.
KW    Hypothetical protein; Helicase; ATP-binding; Complete proteome.
FT    NP_BIND    172    179    ATP (POTENTIAL).
FT    SITE       271    274    DEAQ BOX.
FT    CONFLICT   462    462    D -> S (IN REF. 3).
SQ    SEQUENCE   600 AA;  68873 MW;  75A82D909F4A877A CRC64;

Alignment Scores:
Pred. No.:             31
Score:                 8.00
Percent Similarity:    100.00%
Best Local Similarity: 100.00%
Query Match:           1.43%
DB:                    1

Length:       600
Matches:      8
Conservative: 0
Mismatches:   0
Indels:       0
Gaps:         0

US-09-270-437D-5 (1-1708) x Y018_MYCGE (1-600)

Qy    1172  AATTTGACGGCAGCTGAGGTGGTA 1195
            ||||||||||||||||||||||||
Db    515   AsnLeuThrAlaAlaGluValVal 522


RESULT 11
SG1_BOVIN
ID    SG1_BOVIN    STANDARD;    PRT;    646 AA.
AC    P23389; O02707;
DT    01-NOV-1991 (Rel. 20, Created)
DT    16-OCT-2001 (Rel. 40, Last sequence update)
DT    28-FEB-2003 (Rel. 41, Last annotation update)
DE    Secretogranin I precursor (SgI) (Chromogranin B) (CgB) [Contains: GAWK
DE    peptide; Secretolytin].
GN    CHGB.
OS    Bos taurus (Bovine).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC    Bovidae; Bovinae; Bos.
OX    NCBI_TaxID=9913;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Adrenal chromaffin;
RX    MEDLINE=91223091; PubMed=2025642;
RA    Bauer J.W., Fischer-Colbrie R.;
```

```
RT    "Primary structure of bovine chromogranin B deduced from cDNA
RT    sequence.";
RL    Biochim. Biophys. Acta 1089:124-126(1991).
RN    [2]
RP    SEQUENCE FROM N.A.
RC    TISSUE=Adrenal medulla;
RX    MEDLINE=97282588; PubMed=9136897;
RA    Yoo S.H., Kang Y.K.;
RT    "Identification of the secretory vesicle membrane binding region of
RT    chromogranin B.";
RL    FEBS Lett. 406:259-262(1997).
RN    [3]
RP    SEQUENCE OF 21-646 FROM N.A.
RC    TISSUE=Adrenal medulla;
RA    Grandy D.K., Leduc R., Makam H., Flanagan T., Diliberto E.J.,
RA    Thomas G., Civelli O., Viveros O.H.;
RL    Submitted (OCT-1990) to the EMBL/GenBank/DDBJ databases.
RN    [4]
RP    SEQUENCE OF 634-646.
RC    TISSUE=Adrenal chromaffin;
RX    MEDLINE=95262699; PubMed=7744058;
RA    Strub J.-M., Garcia-Sablone P., Lonning K., Taupenot L., Hubert P.,
RA    van Dorsselaer A., Aunis D., Metz-Boutigue M.-H.;
RT    "Processing of chromogranin B in bovine adrenal medulla.
RT    Identification of secretolytin, the endogenous C-terminal fragment of
RT    residues 614-626 with antibacterial activity.";
RL    Eur. J. Biochem. 229:356-368(1995).
RN    [5]
RP    CHARACTERIZATION OF SECRETOLYTIN.
RX    MEDLINE=96184581; PubMed=8603705;
RA    Strub J.M., Hubert P., Nullans G., Aunis D., Metz-Boutigue M.-H.;
RT    "Antibacterial activity of secretolytin, a chromogranin B-derived
RT    peptide (614-626), is correlated with peptide structure.";
RL    FEBS Lett. 379:273-278(1996).
CC    -!- FUNCTION: Secretogranin I is a neuroendocrine secretory granule
CC        protein, which may be the precursor for other biologically active
CC        peptides. The 16 pairs of basic AA distributed throughout its
CC        sequence may be used as proteolytic cleavage sites.
CC    -!- FUNCTION: Secretolytin has antibacterial activity.
CC    -!- SUBCELLULAR LOCATION: Neuroendocrine and endocrine secretory
CC        granules.
CC    -!- PTM: O-glycosylated (Probable).
CC    -!- SIMILARITY: Belongs to the chromogranin / secretogranin protein
CC        family.
CC    -------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    -------------------------------------------------------------------
DR    EMBL; X55027; CAA38846.1; -.
DR    EMBL; U88551; AAC48720.1; -.
DR    EMBL; X55489; CAA39109.1; -.
DR    PIR; S15901; S15901.
DR    InterPro; IPR001819; Chromogranin_AB.
DR    InterPro; IPR001990; Granin.
DR    Pfam; PF01271; Granin; 1
DR    PRINTS; PR00659; CHROMOGRANIN.
DR    PROSITE; PS00422; GRANINS_1; 1.
DR    PROSITE; PS00423; GRANINS_2; 1.
KW    Sulfation; Cleavage on pair of basic residues; Signal.
FT    SIGNAL       1     20
FT    CHAIN       21    646            SECRETOGRANIN I.
FT    PEPTIDE    418    484            GAWK PEPTIDE.
FT    PEPTIDE    634    646            SECRETOLYTIN.
FT    DISULFID    36     57            BY SIMILARITY.
FT    MOD_RES    158    158            SULFATION (POTENTIAL).
FT    MOD_RES    315    315            SULFATION (BY SIMILARITY).
FT    CONFLICT    64     64            N -> S (IN REF. 1).
FT    CONFLICT    70     70            N -> D (IN REF. 2).
```

```
FT   CONFLICT      93         98       SEAPGL -> FRSPRAS (IN REF. 3).
FT   CONFLICT     181        181       T -> M (IN REF. 2).
FT   CONFLICT     261        261       H -> R (IN REF. 2).
FT   CONFLICT     386        386       P -> R (IN REF. 2).
FT   CONFLICT     481        481       H -> L (IN REF. 3).
FT   CONFLICT     597        597       M -> V (IN REF. 2).
SQ   SEQUENCE   646 AA;  73339 MW;  420DB1178FD9E415 CRC64;

Alignment Scores:
Pred. No.:                           30.8           Length:         646
Score:                                8.00          Matches:          8
Percent Similarity:                 100.00%         Conservative:     0
Best Local Similarity:              100.00%         Mismatches:       0
Query Match:                          1.43%         Indels:           0
DB:                                   1             Gaps:             0

US-09-270-437D-5 (1-1708) x SG1_BOVIN (1-646)

Qy    789 CCAGCAGGTTACTGGGGCTGCTC 812
              ||||||||||||||||||||||
Db      3 ProAlaAlaLeuLeuGlyLeuLeu 10


RESULT 12
BAC1_MOUSE
ID   BAC1_MOUSE       STANDARD;      PRT;      739 AA.
AC   P97302;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Transcription regulator protein BACH1 (BTB and CNC homolog 1).
GN   BACH1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=97042438; PubMed=8887638;
RA   Oyake T., Itoh K., Motohashi H., Hayashi N., Hoshino H., Nishizawa M.,
RA   Yamamoto M., Igarashi K.;
RT   "Bach proteins belong to a novel family of BTB-basic leucine zipper
RT   transcription factors that interact with MafK and regulate
RT   transcription through the NF-E2 site.";
RL   Mol. Cell. Biol. 16:6083-6095(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMRI; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Transcriptional regulator that acts as repressor or
CC       activator. Binds, in-vitro, to NF-E2 binding sites. Play important
CC       roles in coordinating transcription activation and repression by
CC       MAFK.
CC   -!- SUBUNIT: Heterodimer of BACH1 and MAFK.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- TISSUE SPECIFICITY: Ubiquitous.
CC   -!- SIMILARITY: Belongs to the bZIP family. CNC subfamily.
CC   -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D86603; BAA13137.1; -.
DR   EMBL; BC057894; AAH57894.1; -.
DR   HSSP; P34707; 1SKN.
DR   TRANSFAC; T04793; -.
DR   MGD; MGI:894680; Bach1.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0003700; F:transcription factor activity; IDA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.
DR   InterPro; IPR000210; BTB_POZ.
DR   InterPro; IPR008917; Euk_transcr_DNA.
DR   InterPro; IPR002112; Leuzip_Jun.
DR   InterPro; IPR004827; TF_bZIP.
DR   Pfam; PF00651; BTB; 1.
DR   Pfam; PF00170; bZIP; 1.
DR   PRINTS; PR00043; LEUZIPPRJUN.
DR   SMART; SM00338; BRLZ; 1.
DR   SMART; SM00225; BTB; 1.
DR   PROSITE; PS50097; BTB; 1.
DR   PROSITE; PS50217; BZIP; 1.
DR   PROSITE; PS00036; BZIP_BASIC; 1.
KW   Transcription regulation; Activator; Repressor; DNA-binding;
KW   Nuclear protein.
FT   DOMAIN       34        100       BTB.
FT   DNA_BIND    565        580       BASIC MOTIF.
FT   DOMAIN      588        610       LEUCINE-ZIPPER.
SQ   SEQUENCE   739 AA;  81373 MW;  CE2DE606B05F6E32 CRC64;

Alignment Scores:
Pred. No.:                           30.2           Length:         739
Score:                                8.00          Matches:          8
Percent Similarity:                 100.00%         Conservative:     0
Best Local Similarity:              100.00%         Mismatches:       0
Query Match:                          1.44%         Indels:           0
DB:                                   1             Gaps:             0

US-09-270-437D-5 (1-1708) x BAC1_MOUSE (1-739)

Qy    987 ATACGAACTTGGAGTCAGGTGTT 964
              ||||||||||||||||||||||
Db    239 IleArgThrLeuGluSerGlyVal 246


RESULT 13
GYRB_MYXXA
ID   GYRB_MYXXA       STANDARD;      PRT;      815 AA.
AC   O33367;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA gyrase subunit B (EC 5.99.1.3).
GN   GYRB.
OS   Myxococcus xanthus.
OC   Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC   Cystobacterineae; Myxococcaceae; Myxococcus.
OX   NCBI_TaxID=34;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ER-15;
RX   MEDLINE=98304088; PubMed=9639935;
```

RA   Paitan Y., Boulton N., Ron E.Z., Rosenberg E., Orr E.;
RT   "Molecular analysis of the DNA gyrB gene from Myxococcus xanthus.";
RL   Microbiology 144:1641-1647(1998).
CC   -!- FUNCTION: DNA gyrase negatively supercoils closed circular double-
CC   stranded DNA in an ATP-dependent manner and also catalyzes the
CC   interconversion of other topological isomers of double-stranded
CC   DNA rings, including catenanes and knotted rings.
CC   -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC   of double-stranded DNA.
CC   -!- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
CC   breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
CC   enzyme forms an A2B2 tetramer.
CC   -!- SIMILARITY: Belongs to the type II topoisomerase family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AJ000543; CAA04176.1; -.
DR   HSSP; P06982; 1AJ6.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR002288; DNA_gyraseB_C.
DR   InterPro; IPR000565; DNA_GyrB.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR006171; Toprim_dom.
DR   Pfam; PF00204; DNA_gyraseB; 1.
DR   Pfam; PF00986; DNA_gyraseB_C; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF01751; Toprim; 1.
DR   PRINTS; PR00418; TPI2FAMILY.
DR   ProDom; PD149633; DNA_gyraseB_C; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   TIGRFAMs; TIGR01059; gyrB; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Topoisomerase; Isomerase; ATP-binding;
SQ   SEQUENCE   815 AA;   89636 MW;   3862685FBB805B32 CRC64;

Alignment Scores:
Pred. No.:
Score: 29.9
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.43%
DB:

US-09-270-437D-5 (1-1708) x GYRB_MYXXA (1-815)

Qy 1653 GAGGGTGGATCACACCTCAGTGGG 1676
              |||||||||||||||||||||||||
Db   289 GluGlyGlySerHisLeuSerGly 296

RESULT 14
Y018_MYCPN
ID   Y018_MYCPN   STANDARD;   PRT;   1030 AA.
AC   P75093;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical helicase MG018/MG017/MG016 homolog (D12_orf1030).
GN   MPN020 OR MP134.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,

RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
CC   -!- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
CC   -!- SIMILARITY: TO M.GENTALIUM MG016, MG017 AND MG018.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AE000015; AAB95782.1; -.
DR   PIR; S73460, S73460.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR000330; SNF2_N.
DR   InterPro; IPR007527; Znf_SWIM.
DR   Pfam; PF00271; helicase_C; 1.
DR   Pfam; PF00176; SNF2_N; 1.
DR   Pfam; PF04434; SWIM; 1.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
KW   Hypothetical protein; Helicase; ATP-binding; Complete proteome.
FT   NP_BIND   603   610   ATP (POTENTIAL).
FT   SITE   702   705   DEAQ BOX.
SQ   SEQUENCE   1030 AA;   119601 MW;   7272E2B162AF1737 CRC64;

Alignment Scores:
Pred. No.:   29
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.43%
DB:

US-09-270-437D-5 (1-1708) x Y018_MYCPN (1-1030)

Qy 1172 AATTTGACGGCAGCTGAGGTGGTA 1195
              |||||||||||||||||||||||||
Db   945 AsnLeuThrAlaAlaGluValVal 952

RESULT 15
TOP2_ARATH
ID   TOP2_ARATH   STANDARD;   PRT;   1473 AA.
AC   P30182; Q38807;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA topoisomerase II (EC 5.99.1.3).
GN   TOP2 OR AT3G23890 OR F14O13.7.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=95148754; PubMed=7846176;
RA   Xie S., Lam E.;
RT   "Characterization of a DNA Topoisomerase II cDNA from Arabidopsis
RT   thaliana.";
RL   Plant Physiol. 106:1701-1702(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia, and cv. Kas-1;
RX   MEDLINE=95140639; PubMed=7838729;
RA   Xie S., Lam E.;
RT   "Abundance of nuclear DNA topoisomerase II is correlated with

RT   proliferation in Arabidopsis thaliana.";
RL   Nucleic Acids Res. 22:5729-5736(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20363099; PubMed=10907853;
RA   Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT   Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
RT   and BAC clones.";
RL   DNA Res. 7:217-221(2000).
RN   [4]
RP   SEQUENCE OF 751-838 FROM N.A.
RA   Gerhold D., Parsons A., Hadwiger L.A.;
RL   Submitted (XXX-1992) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Control of topological states of DNA by transient
CC   breakage and subsequent rejoining of DNA strands. Topoisomerase II
CC   makes double-strand breaks.
CC   -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC   of double-stranded DNA.
CC   -!- SUBUNIT: Homodimer.
CC   -!- MISCELLANEOUS: Eukaryotic topoisomerase I and II can relax both
CC   negative and positive supercoils, whereas prokaryotic enzymes
CC   relax only negative supercoils.
CC   -!- SIMILARITY: Belongs to the type II topoisomerase family.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; L21015; AAA65448.1; -.
DR   EMBL; M84654; AAA32877.1; -.
DR   EMBL; AP001297; BAB03006.1; -.
DR   EMBL; U12284; AAC48999.1; -.
DR   EMBL; U12285; AAC49000.1; -.
DR   PIR; S53598; S53599.
DR   HSSP; P06786; 1BGW.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR003957; CBFA_NFYB_topis.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002205; DNA_topoisoIV.
DR   Pfam; PF00204; DNA_gyraseB; 1.
DR   Pfam; PF00521; DNA_topoisoIV; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00615; CCAATSUBUNTA.
DR   PRINTS; PR00418; TPI2FAMILY.
DR   ProDom; PD000742; DNA_topoisoIV; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Isomerase; Topoisomerase; DNA-binding; ATP-binding.
FT   NP_BIND   161   166   ATP (POTENTIAL).
FT   ACT_SITE   794   794   DNA CLEAVAGE (BY SIMILARITY).
FT   VARIANT   1213   1213   K -> N (IN STRAIN CV. KAS-1).
FT   VARIANT   1245   1245   A -> G (IN STRAIN CV. KAS-1).
FT   VARIANT   1299   1299   E -> G (IN STRAIN CV. KAS-1).
SQ   SEQUENCE   1473 AA; 164106 MW; 00B6C4836E381403 CRC64;

Alignment Scores:
Pred. No.:   Length: 1473
Score: 27.7   Matches: 8
Percent Similarity: 8.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 1.44%   Gaps: 0
                         1

US-09-270-437D-5 (1-1708) x TOP2_ARATH (1-1473)

Qy   538 CTTCTTCAGTTCCGTCCTTCCTT 515
           |||||||||||||||||||||||
Db   561 LeuLeuGlnValProSerPheLeu 568

Search completed: July 13, 2004, 12:31:31
Job time : 31.5 secs

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:    July 13, 2004, 12:12:42 ; Search time 87 Seconds
                                      (without alignments)
                                      12247.273 Million cell updates/sec

Title:         US-09-270-437D-5
Perfect score: 3110
Sequence:      1 aggacgctgcgcaccgcc........atttcctcaggtttaaaa 1708

Scoring table:  BLOSUM62
                Xgapop  10.0 ,  Xgapext  0.5
                Ygapop  10.0 ,  Ygapext  0.5
                Fgapop   6.0 ,  Fgapext  7.0
                Delop    6.0 ,  Delext   7.0

Searched:      1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters:       2559352

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_13072004_121924_9724/app_query.fasta_1.1863
-DB=Published Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09270437 @CGN 1_1.130 @runat_13072004_121924_9724
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  Published Applications_AA:*
   1:  /cgn2_6/ptcdata/1/pubpaa/US07_PUBCOMB.pep:*
   2:  /cgn2_6/ptcdata/1/pubpaa/1/pubpaa/PCT_NEW_PUB.pep:*
   3:  /cgn2_6/ptcdata/1/pubpaa/US06_NEW_PUB.pep:*
   4:  /cgn2_6/ptcdata/1/pubpaa/US06_PUBCOMB.pep:*
   5:  /cgn2_6/ptcdata/1/pubpaa/US07_NEW_PUB.pep:*
   6:  /cgn2_6/ptcdata/1/pubpaa/1/pubpaa/PCTUS_PUBCOMB.pep:*
   7:  /cgn2_6/ptcdata/1/pubpaa/US08_NEW_PUB.pep:*
   8:  /cgn2_6/ptcdata/1/pubpaa/US08_PUBCOMB.pep:*
   9:  /cgn2_6/ptcdata/1/pubpaa/US09A_PUBCOMB.pep:*
  10:  /cgn2_6/ptcdata/1/pubpaa/US09B_PUBCOMB.pep:*
  11:  /cgn2_6/ptcdata/1/pubpaa/US09C_PUBCOMB.pep:*
  12:  /cgn2_6/ptcdata/1/pubpaa/US09_NEW_PUB.pep:*
  13:  /cgn2_6/ptcdata/1/pubpaa/US10A_PUBCOMB.pep:*
  14:  /cgn2_6/ptcdata/1/pubpaa/US10B_PUBCOMB.pep:*
  15:  /cgn2_6/ptcdata/1/pubpaa/US10C_PUBCOMB.pep:*
  16:  /cgn2_6/ptcdata/1/pubpaa/US10_NEW_PUB.pep:*
  17:  /cgn2_6/ptcdata/1/pubpaa/US60_NEW_PUB.pep:*
  18:  /cgn2_6/ptcdata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
Result         %
No.    Score  Query  Match Length DB  ID                          Description
-----  -----  -----  ----- ------ --  --                          -----------
```

1	2226	71.6	577	15	US-10-313-986-500	Sequence 500, App
2	2208	71.0	577	9	US-09-873-637-2	Sequence 2, Appli
3	1639	52.7	579	9	US-09-735-705-148	Sequence 348, App
4	1639	52.7	579	9	US-09-850-716A-348	Sequence 348, App
5	1639	52.7	579	9	US-09-897-778-348	Sequence 348, App
6	1639	52.7	579	9	US-09-897-778-446	Sequence 446, App
7	1639	52.7	579	9	US-09-897-778-449	Sequence 449, App
8	1639	52.7	579	12	US-10-007-700-348	Sequence 348, App
9	1639	52.7	579	12	US-10-007-700-446	Sequence 446, App
10	1639	52.7	579	12	US-10-007-700-449	Sequence 449, App
11	1639	52.7	579	14	US-10-117-982-348	Sequence 348, App
12	1639	52.7	579	14	US-10-117-982-446	Sequence 446, App
13	1639	52.7	579	14	US-10-117-982-449	Sequence 449, App
14	1639	52.7	579	14	US-10-117-982-480	Sequence 480, App
15	1639	52.7	579	15	US-10-313-986-348	Sequence 348, App
16	1639	52.7	579	15	US-10-313-986-446	Sequence 446, App
17	1639	52.7	579	15	US-10-313-986-449	Sequence 449, App
18	1639	52.7	579	15	US-10-313-986-480	Sequence 480, App
19	1639	52.7	586	9	US-09-850-716A-427	Sequence 427, App
20	1639	52.7	586	9	US-09-897-778-427	Sequence 427, App
21	1639	52.7	586	12	US-10-007-700-427	Sequence 427, App
22	1639	52.7	586	14	US-10-117-982-427	Sequence 427, App
23	1639	52.7	586	15	US-10-313-986-427	Sequence 427, App
24	1639	52.7	589	15	US-10-313-986-486	Sequence 486, App
25	1637	52.6	579	9	US-09-735-705-176	Sequence 176, App
26	1637	52.6	579	9	US-09-850-716A-176	Sequence 176, App
27	1637	52.6	579	9	US-09-897-778-176	Sequence 176, App
28	1637	52.6	579	10	US-09-466-396A-176	Sequence 176, App
29	1637	52.6	579	12	US-10-007-700-176	Sequence 176, App
30	1637	52.6	579	14	US-10-117-982-176	Sequence 176, App
31	1637	52.6	579	15	US-10-313-986-176	Sequence 176, App
32	1635	52.6	579	14	US-10-117-982-484	Sequence 484, App
33	1635	52.6	579	15	US-10-313-986-484	Sequence 484, App
34	1560	50.2	422	16	US-10-408-765A-2088	Sequence 2088, Ap
35	1530.5	49.2	620	9	US-09-764-864-1116	Sequence 1116, Ap
36	1524.5	49.0	587	15	US-10-313-986-501	Sequence 501, App
37	1464	47.1	556	14	US-10-097-340-147	Sequence 147, App
38	1464	47.1	556	16	US-10-648-593-182	Sequence 182, App
39	1355.5	43.6	555	15	US-10-262-445-40	Sequence 40, Appl
40	1258	40.5	261	9	US-09-764-864-1114	Sequence 1114, Ap
41	1228	39.5	250	9	US-09-764-864-1532	Sequence 1532, Ap
42	653	21.0	171	9	US-09-764-864-1119	Sequence 1119, Ap
43	510	16.4	192	9	US-09-764-864-1117	Sequence 1117, Ap
44	502	16.1	171	9	US-09-764-864-1536	Sequence 1536, Ap
45	361	11.6	81	14	US-10-117-982-476	Sequence 476, App

ALIGNMENTS

```
RESULT 1
US-10-313-986-500
; Sequence 500, Application US/10313986
; Publication No. US2003023620A1
; GENERAL INFORMATION:
;  APPLICANT:  Foy, Teresa M.
;  APPLICANT:  McNabb, Andria
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Wang, Tongtong
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C19
;  CURRENT APPLICATION NUMBER: US/10/313,986
;  CURRENT FILING DATE:  2002-12-04
;  NUMBER OF SEQ ID NOS:  560
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 500
;  LENGTH: 577
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-10-313-986-500

  Alignment Scores:
```

```
Pred. No.:                          577
Score:              2.18e-168       Length:             577
Percent Similarity:  99.33%         Matches:            2226.00
Best Local Similarity: 99.33%       Conservative:       0
Query Match:         71.58%         Mismatches:         1
DB:                                 Indels:             2
                                    Gaps:               1
```

US-09-270-437D-5 (1-1708) x US-10-313-986-500 (1-577)

```
Qy     32 CGGGGAGCCATCATGAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC 91
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 ArgGlnAlaIleMetLysLeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSer 152

Qy     92 TACATCCCCGATGAGCAGATAGCACAGGAGCCTGAGAATGGCGGCCGAGGGGCTTTGGC 151
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 TyrIleProAspGluGlnIleAlaGlnGluProGluAsnGlyGlyArgGlyGlyPheGly 172

Qy    152 TCTCGGGGTCAGCCCCGCCAGGGCTCACCTGTGCAGCGGGGCCCCAGCCAAGCAGCAG 211
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 SerArgGlyGlnProArgGlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGln 192

Qy    212 CAAGTGGACATCCCCTTCGGTCCTGTGCCCACCCAGTATGTGGGTGCCATTATTGGC 271
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 GlnValAspIleProLeuArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGly 212

Qy    272 AAGGAGGGGCCACCATCCCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCAT 331
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 LysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHis 232

Qy    332 AGGAAGGAGACGCAGTCGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGC 391
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    233 ArgLysGluAsnAlaGlyAlaAlaGlyLysAlaIleSerValHisSerThrProGluGly 252

Qy    392 TGCTCCCCGCCTGTAAGGATGATCTTGGAGATTATGCATAAAGAGAGCTAAGGACACCAAA 451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 CysSerSerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLys 272

Qy    452 ACGGCTGACGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATT 511
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 ThrAlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIle 292

Qy    512 GGGAAGGAGGGACGGAACCTGAAGAAGTAGAGCAAGATACCGACAGCAAAAATCACCATC 571
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    293 GlyLysGluGlyArgAsnLeuLysLysValGluGlnAspThrGluThrLysIleThrIle 312

Qy    572 TCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGAGACCATCACTGTGAAGGGGGCC 631
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 SerSerLeuGlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAla 332

Qy    632 ATCGAGAATGTTGCAGGGGCCGAGCAGAGAAATAATGAAGAAAGTTCGGGAGGCCTATGAG 691
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 IleGluAsnCysCysArgAlaGluGlnGlnLysMetLysLysValAlaArgGluAlaTyrGlu 352

Qy    692 AATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGCCTGAACCTGGCTGCT 745
          |||||||||||||||||            ||||||||||||||||||||||||||||
Db    353 AsnAspValAlaAlaMetSerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAla 372

Qy    746 GTAGGTCTTTTCCCACTTCATCCAGCGCAGTCCCGCGCCTTCCCGAGCGCGTTACTGGG 805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 ValGlyLeuPheProLeuHisProAlaGlnSerArgAlaSerAlaValProProProProSerValThrGly 392

Qy    806 GCTGCTCCCTATAGCTCCTTTATGCAGAGCTCCCAGAGCAGGAGATGGTGCAGGTGTTTATC 865
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    393 AlaAlaProTyrSerSerPheMetGlnSerSerGlnSerArgArgProGluGlnAlaProGlyGluGlnMetValGlnValPheIle 412

Qy    866 CCCGCCAGCAGTGGGCGGCCATCATCGGCAAGAGGGGCAGCACATCAAACAGCTCTCC 925
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    413 ProAlaGlnAlaValAlaGlyIleAlaIleIleGlyLysLysGlyGlnHisIleIleGlnLeuSer 432

Qy    926 CGGTTTGCCAGCCGCGCTCCAATCAAGATTGCACCAGGAAAACACTGACTCACAGTTCGT 985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    433 ArgPheAlaSerAlaSerIleLysLysIleAlaProGlyLysHisGln ...ArgLysValArg 452
```

```
Qy    986 ATGGTTATCATCACTGGACCGCCAGAGGCCCAATTCAGGCTCAGGAAGAATCTATGGC 1045
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    453 MetValIleIleThrGlyProProGluAlaGlnPheLysLeuArgLysAsnLeuTyrGly 472

Qy   1046 AAACTCAAGAGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGCTGAGCTGGAGACCCACATA 1105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    473 LysLeuLysGluGluLeuLeuTrpSerGlnGlyGlySerLeuLeuGluThrHisIle 492

Qy   1106 CGTGTGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGTGGAAAAACGGTGAACGAG 1165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    493 ArgValProAlaSerAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGlu 512

Qy   1166 TTGCAGAATTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATCGAGAAC 1225
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    513 LeuGlnAsnLeuThrAlaAlaGluValValValValProArgAspProArgAspGluAsn 532

Qy   1226 GACCAGGTCATCGTGAAAATCATCGGACATTTCATGCAGGTCAGATGGCTCACGGAAG 1285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    533 AspGlnValIleValLysIleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLys 552

Qy   1286 ATCCGAGACATCCTGCCCAGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACCAGGCC 1345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    553 IleArgAspIleLeuProSerLeuArgLeuGlnProGlnAsnGlnGlnAlaGlnGlnSerAsnGlnAla 572

Qy   1346 CAGGCACGGAGGAAG 1360
          |||||||||||||||
Db    573 GlnAlaArgArgLys 577
```

RESULT 2
US-09-873-637-2
; Sequence 2, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-2

Alignment Scores:
```
Pred. No.:                          577
Score:              5.91e-167       Length:             577
Percent Similarity:  98.65%         Matches:            2208.00
Best Local Similarity: 98.43%       Conservative:       1
Query Match:         71.00%         Mismatches:         4
DB:                                 Indels:             2
                                    Gaps:               1
```

US-09-270-437D-5 (1-1708) x US-09-873-637-2 (1-577)

```
Qy     32 CGGGGAGCCATCATGAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC 91
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 ArgGlnAlaIleMetLysLeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSer 152

Qy     92 TACATCCCCGATGAGCAGATAGCACAGGAGCCTGAGAATGGCGGCGGGGGCTTTGGC 151
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 TyrIleProAspGluGlnIleThrGlnGlyProGluAsnGlyArgArgGlyGlyPheGly 172

Qy    152 TCTCGGGGTCAGCCCCGGCTCACCTGTGCAGCGGGGCCCCAGCCAAGCAGCAG 211
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 SerArgGlyGlnProArgGlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGln 192

Qy    212 CAAGTGGACATCCCCTTGGGTCCTGTGCCCACCCAGTATGTGGGTGCCATTATTGGC 271
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 GlnValAspIleProLeuGlyLeuLeuValProThrGlnTyrValGlyAlaIleIleGly 212

Qy    272 AAGGAGGGGCCACCATCCCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCAT 331
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 LysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHis 232
```

```
Db    573 GlnAlaArgArgLys 577
          |||||||||||||
          GlnAlaArgArgLys 577

RESULT 3
US-09-735-705-348
; Sequence 348, Application US/09735705
; Patent No. US2002052329A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT: Fan, Liqun
;  APPLICANT: Kalos, Michael D.
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT: Hosken, Nancy
;  APPLICANT: Fanger, Gary R.
;  APPLICANT: Li, Samuel X.
;  APPLICANT: Wang, Aijun
;  APPLICANT: Skeiky, Yasir A.W.
;  APPLICANT: Henderson, Robert A.
;  APPLICANT: McNeill, Patricia D.
;  APPLICANT: Fanger, Neil
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C14
;  CURRENT APPLICATION NUMBER: US/09/735,705
;  CURRENT FILING DATE: 2000-12-12
;  NUMBER OF SEQ ID NOS: 419
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
;  LENGTH: 579
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-735-705-348

Alignment Scores:
Pred. No.:              1.26e-121        Length:        579
Score:                  1639.00          Matches:       336
Percent Similarity:     83.89%           Conservative:  44
Best Local Similarity:  74.17%           Mismatches:    57
Query Match:            52.70%           Indels:        16
DB:                     9                Gaps:          9

US-09-270-437D-5 (1-1708) x US-09-735-705-348 (1-579)
```


Sequence alignment data (left column, Qy/Db pairs with nucleotide and amino acid sequences numbered 272–1360).

```
;   CURRENT APPLICATION NUMBER: US/09/850,716A
;   CURRENT FILING DATE: 2001-05-07
;   NUMBER OF SEQ ID NOS: 440
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-850-716A-348

Alignment Scores:
Pred. No.:                                Length:        579
Score:                    1.26e-121       Matches:       336
Percent Similarity:       1639.00         Conservative:   44
Best Local Similarity:    83.89%          Mismatches:     57
Query Match:              74.17%          Indels:         16
DB:                       52.70%          Gaps:            9
                          9

US-09-270-437D-5 (1-1708)  x  US-09-850-716A-348  (1-579)
```

```
Qy      32  CGGGGAGCCATCATGAAGTCAATGAGCTCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC   91
            |||    |||  ||||||||||  :::::||  ||||||||||||||||||||  :::
Db     133  ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuLeuGluAsnPheThrLeuLysLysValAla  152

Qy      92  TACATCCCGATGACCAGATAGCA---CAGGGACCTGAGAATGGGCGCCAGGG-----  142
            |||||||||||||||||| |||       |||||||||  |||||||||||
Db     153  TyrIleProAspGluThrAlaAlaGlnGlnSerProLeuGlnGlnProArgGlyIleProArgArg  172

Qy     143  GGCTTTGGCTCTCGGGGTCAGCCCCAGGGGTCACCGTGGGGGGGGCCCCAGCC  202
            :::  |||||  ||||||||||||||||||||       ||||||||||
Db     173  GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro---------GlySerValSer  189

Qy     203  AAGCAGCAGCAAGTGACATCCCCTTCGGCTCCTGGTGCCCAGTATGTGGGTGCC  262
            ||||:::        |||||||||||||||||  |||||||||||:::::||||
Db     190  LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla  209

Qy     263  ATTATTGGCAAGGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATA  322
            ||||||  ||||||||  ||||||||||||||||||||||  ||||||||||
Db     210  IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle  229

Qy     323  GACGTGCATAGGAAGGAAGGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC  382
            |||  |||||||||||||||||||||||||||||||||||||||
Db     230  AspValHisArgLysGluGluGluThrGlnValGlnLeuLysSerHisIleThrLeuSerThr  249

Qy     383  CCTGAGGGGTGCTCCTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG  442
            |||:::     |||||||||||||||||||||||  ||||||:::|||||||||
Db     250  ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln  269

Qy     443  GACACCAAAACGGCTGACGAGGTTCCCCATAATAACTTTGTAGGG  502
            |||  |||||||||||||||||||||||||||||||||||||
Db     270  AspIleLysPheThrGluLlleProLeuLysIleLeuAlaHisAsnAsnPheValGly  289

Qy     503  CGTCTCATTGGCAAGGAAGGAGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAA  562
            ||||||||||||||||||||||||||||||||||||||||  |||||||||
Db     290  ArgLeuIleGlyLysGluGlyAspGlyThrArgAsnLeuLysLysIleLysIleGluGlnGlnThrArgLys  309

Qy     563  ATCACCATCTCCTGTTGCAAGACCTTACACCCTTACACACCTGAGAGAGGACCATCACTGTG  622
            |||||||||  :::     ||||||||||||||||||||||:::|||||
Db     310  IleThrIleSerProLeuGlnLeuThrLeuTyrAsnProGluArgThrIleThrVal  329

Qy     623  AAGGGGGCCATCATGAAGAATTGTTGCAGGGCCGAGCGAGGAAATATGAAGAAGAAGTTGGGAG  682
            |||    |||  :::||::|||||  ::::  :::|||||||||:::
Db     330  LysGlyAsnValGluThrCysAlaLysSerAlaGluGluGluIleIleMetLysLysIleArgGlu  349

Qy     683  GCCTATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCCTGAAC  736
            ::::|||||::::::     ||||||       :::|||||||||||||||||||
Db     350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnLeuAlaHisLeuIleProGlyLeuAsn  369

Qy     737  CTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCCAGTCCCGCCG------CCT  787
            |||     ||||||  :::  :::||:::||||||||||||||||
Db     370  LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyPro  389

Qy     788  CCCAGCAGCCGTTACTGGGGCTGCTCCCTATAGTCTCCCTTTATGCAGGCTCCCGAGCAGGAG  847
```

```
Db     270  AspIleLysPheThrGluLlleProLeuLysIleLeuAlaHisAsnAsnPheValGly  289

Qy     503  CGTCTCATTGGCAAGGAAGGAGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAA  562

Db     290  ArgLeuIleGlyLysGluGlyAspGlyThrArgAsnLeuLysLysIleLysIleGluGlnGlnThrArgLys  309

Qy     563  ATCACCATCTCCTGTTGCAAGACCTTACACCCTTACACACCTGAGAGAGGACCATCACTGTG  622

Db     310  IleThrIleSerProLeuGlnLeuThrLeuTyrAsnProGluArgThrIleThrVal  329

Qy     623  AAGGGGGCCATCATGAAGAATTGTTGCAGGGCCGAGCGAGGAAATATGAAGAAGAAGTTGGGAG  682

Db     330  LysGlyAsnValGluThrCysAlaLysSerAlaGluGluGluIleIleMetLysLysIleArgGlu  349

Qy     683  GCCTATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCCTGAAC  736

Db     350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnLeuAlaHisLeuIleProGlyLeuAsn  369

Qy     737  CTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCCAGTCCCGCCG------CCT  787

Db     370  LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyPro  389

Qy     788  CCCAGCAGCCGTTACTGGGGCTGCTCCCTATAGTCTCCCTTTATGCAGGCTCCCGAGCAGGAG  847
Db     390  ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu  406

Qy     848  ATGGTGAGGTGTTTATCCCCGCCAGGCAGTGGGCGCCATCATCCGCAAGAAGGGGGCAG  907
Db     407  ThrValHisLeuPheIleProAlaLeuSerValGlyIleAlaIleIleGlyLysLysGlnGlyGln  426

Qy     908  CACATCAAACAGCTCCGTTGCCAGCGGCCTCCATCAAGATTGCACCACCCGAAACA  967
Db     427  HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysLysIleAlaProAlaGluAla  446

Qy     968  CCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCCAGAGGCCAATTCAAGGCT  1027
Db     447  ProAspSerLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla  466

Qy    1028  CAGGGGAAGAATTCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG  1087
Db     467  GlnGlyLysAsnSerMetAlaAsnSerArgArgArgAsnPhePheGlyProLysGluGluVal  486

Qy    1088  AAGCTGGAGACCCCACATACGTGTGCCAGCATCAGCAGCTGCGCGGGTCATTGCCAAAGGT  1147
Db     487  LysLeuGluThrProHisThrCysAlaSerIleSerSerCysAlaGlyHisCysGlnArg  506

Qy    1148  GGAAAAACGGTCGAACGAGTTCAGAGATTTGACGGCGAAGAGTTCAAGGAGGTACCAAGAGAC  1207
Db     507  GlyLysThrValAsnGluLeuGlnAsnLeuAsnLeuSerSerAlaGluValValValProArgAsp  526

Qy    1208  CAGACCCCTGATGAGACAACGACCAGTCATCGTGAAAATCATCGGACATTTCTATGCCAGT  1267
Db     527  GlnThrProAspGluThrThrThrSerHisArgGluAsnHisArgThrPheLeuCysGln  546

Qy    1268  CAGATGGCTCAACGGAAGATCCGAGAGATCCTGGCCCAGGTTAAGCAG--CAGCATCAG  1324
Db     547  GlnMetAlaGlnArgLysIleArgGluIleLeuAlaGlnValLysGlnGlnHisGlnGln  566

Qy    1325  AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGGAAG  1360
Db     567  LysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys  579
```

```
RESULT 4
US-09-850-716A-348
;  Sequence 348, Application US/09850716A
;  Patent No. US2002011513941
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
```

RESULT 5
US-09-897-778-348
; Sequence 348, Application US/09897778
; Patent No. US2002014714 3A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 2101 21.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-348

Alignment Scores:
Pred. No.:
Score: 1639.00
Percent Similarity: 83.89%
Best Local Similarity: 74.17%
Query Match: 52.70%
DB:

Length: 579
Matches: 336
Conservative: 44
Mismatches: 57
Indels: 16
Gaps: 9

US-09-270-437D-5 (1-1708) x US-09-897-778-348 (1-579)

```
Db   467  GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486

Qy  1088  AAGCTGGAGACCCACATACGTGTGCCAGTTGCAGGAACCATGCCCTGGGTCATTGGCAAAGGT 1147
          |||||||||||||||||||||||| |||:::  |||||||||||||||||:::||||||||||
Db   487  LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyValIleGlyLysLysGly 506

Qy  1148  GGAAAAACGTGAACGAGTTGCAGAATTGACGGCAGTGAGGTGGTAGTACCAAGAGAC 1207
          |||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db   507  GlyLysThrValAsnGluLeuGlnAsnLeuSerAlaGluValValProArgAsp 526

Qy  1208  CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCCGACATTTCTATGCCAGT 1267
          |||||||||||||||||||||||||:::|||||||||:::||||||||||||||
Db   527  GlnThrProAspGluAsnAspGlnValIleValLysIleIleThrGlyHisPheTyrAlaCys 546

Qy  1268  CAGATGGCTCAACGGAAGATCCGAGACATCCGGCCCAGGTTAAGCAG---CAGCATCAG 1324
          ||:|||||||||||||||||||||||||:::||||||||||||||||
Db   547  GlnValAlaGlnArgLysIleGluGlnIleLeuThrGlnValLysGlnGlnHisGlnGlnGln 566

Qy  1325  AAGGGA--CAGAGTAACCAGGCCCAGGCACGGAGGAAG 1360
          ||||||  |||:|||||||||||||
Db   567  LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579
```

RESULT 6
US-09-897-778-446
; Sequence 446, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Marnerakis, Margarita
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Vedvick, Thomas S.
;   APPLICANT: Carter, Darrick
;   APPLICANT: Watanabe, Yoshihiro
;   APPLICANT: Henderson, Robert A.
;   APPLICANT: Peckham, David W.
;   APPLICANT: Fanger, Neil
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C16
;   CURRENT APPLICATION NUMBER: US/09/897,778
;   CURRENT FILING DATE: 2001-06-28
;   NUMBER OF SEQ ID NOS: 467
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 446
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-897-778-446

```
Alignment Scores:
Pred. No.:                          579
Score:              1.26e-121       Length:     579
Percent Similarity: 1639.00         Matches:    336
Best Local Similarity: 83.89%       Conservative: 44
Query Match:        74.17%          Mismatches: 57
                    52.70%          Indels:     16
DB:                 9               Gaps:       9
```

US-09-270-437D-5 (1-1708) x US-09-897-778-446 (1-579)

```
Qy    32  CGGGGGAGCCATCATGAAGCTGAATGGCACCAGTTGAGAACCATGCCCTGAAGGTCTCC 91
          |||  |||  ||||:::  |||||||||||||||||||||:::||
Db   133  ArgGlnAlaLeuAspLysPheGlnAsnGlyAsnGlyLeuGluAsnPheThrLeuLysValAla 152

Qy    92  TACATCCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGCGGCGCCGAGGG------ 142
          |||||||||||||||||||||    |||||||||||||||||||||
Db   153  TyrIleProAspGluThrAlaAlaAlaGlnIleAlaAlaGlnAsnProLeuGlnArgGlyArgArg 172

Qy   143  GGCTTGGCTCCTCGGGTCAGCCTCACCGTGGCGCCCAGGGCCCCAGCC 202
          ||||||||||||:::||||||||||||||||
Db   173  GlyLeuGlyValGlnArgGlySerArgGlnGlySerPro-------GlySerValSer 189

Qy   203  AAGCAGCAGCAAGTGGACATCCCCTTCGGCTCCTGTGTGCCACCCAGTATGTGGGTGCC 262
          |||:::  ||||||||||||||||:::  ||||||||||
```

```
Db   190  LysGlnLysProCysAspTyrProLeuArgLeuLeuLeuValProThrGlnPheValGlyAla 209

Qy   263  ATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAAGATA 322
          ||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db   210  IleIleIleGlyLysGluGlyAlaAlaThrIleArgAsnIleThrLysGlnSerLysIle 229

Qy   323  GACGTCCATAGGAAGGAGACGACGCAGTGCAGCTCGAAAAAGCCATCAGTGTGCACTCCACC 382
          ||||||||||||||||:::|||||||||||||||||||||
Db   230  AspValHisArgLysGluThrThrGlnAlaAlaArgLysAlaIleSerValHisSerThr 249

Qy   383  CCTGAGGGCTGCTCTCCGGTTGTAAGATGATC------TTGGAGATTATGCATAAAGAGGCTAAG 442
          ||||||||||||||||||||||||:::||
Db   250  ProGluGlyThrSerAlaAlaCysSerLeuLeuMetHisLysGluAlaGln 269

Qy   443  GACACCAAAACGGTGACGAGTTCCCGTGAGATCCTGGCCCATAATAACTTTGTAGGG 502
          ||||||||||||||||||||:::||||||||||||||
Db   270  AspIleLysPheThrGluGlnLeuProLeuLysIleLeuAlaHisAsnAsnPheValGly 289

Qy   503  CGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAA 562
          ||||||||||||||||||:::||||||||||||||
Db   290  ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysValGluGlnAspThrAspThrLys 309

Qy   563  ATCACCATCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGGACCATCACTGTG 622
          ||||||||||||||||||||:::|||||||||||||
Db   310  IleThrIleSerProLeuGlnLeuThrLeuTyrAsnProGluAspHisHisCysVal 329

Qy   623  AAGGGGGCCATCGAGAATTGTTGCAGGCAGGAGAAATAATGAAGAAAGTTCGGGAG 682
          |||||||||||||||||||||||:::|||||||||||||
Db   330  LysGlyAlaIleGluAsnCysCysArgGlnGluLysIleMetLysLysValArgGlu 349

Qy   683  GCCTATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCCTGAAC 736
          |||||||||||||||||||||    |||||||||||||||||
Db   350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

Qy   737  CTGGCTGCTCGTAGTCTTTTCCCCAGCTTCATCCAGCGCAGTCCCCCG-------CCT 787
          ||||||||||||||||||:::      :::||
Db   370  LeuAlaAsnAlaLeuGlyLeuPheProThrSerGlyMetProProProThrGlyPro 389

Qy   788  CCCAGCAGCGTTACTGGGGCTGCTCCTATAGCTCCTTTATGCAGGCTCCCAGCAGGAG 847
          |||||||||||||||||||||:::      :::|||||||||||
Db   390  ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu 406

Qy   848  ATGGTGCAGGTGTTTATCCCGCCAGGCAGTGGGCGCCATCATGGCAAGGAGGAGGGCAG 907
          :::|||||||||||||  :::||||||||
Db   407  ThrValHisLeuPheIleProIleAlaLeuSerValGlyAlaIleIleGlyLysLysGlyGln 426

Qy   908  CACATCAAACAGCTCTCCCGGTTGCAGCGCGTCCATCAAGATTGACCACCCGAAACA 967
          |||||||||||:::      ||||||||||||||
Db   427  HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleAlaProAlaAlaGluAlaA 446

Qy   968  CCTGACTCCAAGTTCGTATGGTTATCATCACTGGACCCGCCAGAGGCCCAATTCAAGGCT 1027
          ||||||||||||||||||||:::      |||||||||||||||
Db   447  ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysAla 466

Qy  1028  CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGACTCTTGGTCCCAAGGAGGAAGTG 1087
          |||||||||||||||:::|||||||||||||||
Db   467  GlnGlyArgIleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluVal 486

Qy  1088  AAGCTGGAGACCCACATACGTGTGCCAGTTGCAGCAGTGGCCGGGTCATTGGCAAAGGT 1147
          |||||||||||||||||||||||| |||:::  |||||||||||||||||
Db   487  LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyValIleGlyLysLysGly 506

Qy  1148  GGAAAAACGTGAACGAGTTGCAGAATTGACGGCAGTGAGGTGGTAGTACCAAGAGAC 1207
          |||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db   507  GlyLysThrValAsnGluLeuGlnAsnLeuSerAlaGluValValProArgAsp 526

Qy  1208  CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCCGACATTTCTATGCCAGT 1267
          ||||||||||||||||||||||:::|||||||||:::|||||||||||||
Db   527  GlnThrProAspGluAsnAspGlnValIleValLysIleIleThrGlyHisPheTyrAlaCys 546

Qy  1268  CAGATGGCTCAACGGAAGATCCGAGACATCCGGCCCAGGTTAAGCAG---CAGCATCAG 1324
          ||:||||||||||||||||||||||:::||||||||||||
Db   547  GlnValAlaGlnArgLysIleGluGlnIleLeuThrGlnValLysGlnGlnHisGlnGlnGln 566
```

Qy   1325 AAGGGA---CAGAGTAACCAGGCCACGGAGGAAG 1360
          ||||| |||||||||||||||||:|||||||||
Db   567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 7
US-09-897-778-449
; Sequence 449, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Marnerakis, Margarita
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Henderson, Robert A.
APPLICANT: Peckham, David W.
APPLICANT: Fanger, Neil
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C16
CURRENT APPLICATION NUMBER: US/09/897,778
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 467
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 449
LENGTH: 579
TYPE: PRT
ORGANISM: Homo sapiens
US-09-897-778-449

Alignment Scores:
Pred. No.:   1.26e-121   Length: 579
Score: 1639.00   Matches: 336
Percent Similarity: 83.89%   Conservative: 44
Best Local Similarity: 74.17%   Mismatches: 57
Query Match: 52.70%   Indels: 16
DB: Gaps: 9

US-09-270-437D-5 (1-1708) x US-09-897-778-449 (1-579)

Qy   32 CGGGGAGCCACATGAGCTGAATGGCCACCAGTTGAGAACCATGCCTGAAGGTCTCC 91
          ||||||| |||:::| |||||||||||||||| |||||||||||:::
Db   133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnHisLeuGluAsnPheThrLeuLysValAla 152

Qy   92 TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGCGGCCGAGGG------ 142
          |||||||| |||||| ||||||| |||||||||||||||||||||
Db   153 TyrIleProAspGluThrAlaAlaGlnLeuAlaGlnLeuAlaAsnProLeuGlnProLeuArgGlyArgArg 172

Qy   143 GGCTTTGGCTCTCGGGGTCAGCCCCGCCAGGGTCACCTGTGCAGCGGGGGCCCCAGCC 202
          ||||| ||| |||||||||||::: |||||||||| ::: ::::
Db   173 GlyLeuGlyGlnArgGlyGlySerArgGlnGlySerPro--------GlySerValSer 189

Qy   203 AAGCAGCAGAAGTGGACATCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCC 262
          |||||||||||:::|||||||:::||||| |||||||||||:::
Db   190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209

Qy   263 ATTATTGGCAAGAAGGGGGGCCACCACCGCCAACATCACAAAACAGACCCAGTCCAAGATA 322
          |||||||||||::: ||||||||||||| ||||||||| ||||||||||||
Db   210 IleIleGlyLysLysGlyGluGlyIleAlaThrIleArgAsnIleThrLysGlnSerLysLysIle 229

Qy   323 GACGTGCATAGGAGGGAAGAAGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC 382
          |||||||| ||||||| ||||| |||| |||||||:::::: |||||
Db   230 AspValHisArgArgLysGluGluAsnAlaGlyAlaAlaGluLysSerIleSerIleLeuSerThr 249

Qy   383 CCTGAGGGGTGCTCCTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGAGGCTAAG 442
          ||||||||||| ||||||||||||| ::: |||:::|||||||||||||
Db   250 ProGluGlyThrSerAlaAlaCysLysMetIleLeuGluIleMetHisLysGluArgGln 269

Qy   443 GACACCAAAACGGCTGACGAGGTTCCCTGAAGATCCTGCCCATAATAACTTTGTAGGG 502
          ||||||| ||| ||||| :::|||||||||||||:::
Db   270 AspIleLysPheThrGluGluLeuIleProAlaLeuLysIleLeuAlaHisAsnAsnPheValGly 289

Qy   503 CGTCTCATTGGCAAGAAGAGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAA 562
          |||||||||||||||||||||:::||||||:::|||||||||
Db   290 ArgLeuIleGlyGluGlyGlyArgAsnLeuLysLysGluGlnIleAspThrGluThrLys 309

Qy   563 ATCACCATCTCCTCGTTGCAAGACTTACCCTTACAACCCTGAGAGGACCATCACTGTG 622
          |||||||||||::: |||||||||||||||:::||||||||||||
Db   310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

Qy   623 AAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGAAATAATGAGAAAGTTCGGGAG 682
          ||||| :::|||||||||:::|||||||||:::|||||||||
Db   330 LysGlyAsnValGluThrCysAlaLysAlaGluGlnGluIleMetLysLysIleArgGlu 349

Qy   683 GCCTATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCTGGCCTGAAC 736
          ||||||||||||||||||||||||||||| :::|||||||||||||||
Db   350 SerTyrGluAsnAspAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

Qy   737 CTGGCCTGCTGTAGGTCTTTCCCAGTTCATCCAGGCAGTCCCGCG-----CCT 787
          ||||||| :::||| :::||||||||||||
Db   370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389

Qy   788 CCCAGCAGCGTTACTGGGGCTGCTCCTATAGCTCCTTTATGCAGGTCCGAGCAGGAG 847
          |||||| :::||| || ------ProProTyrProGlnPheGluGlnSer---GluThrGlu
Db   390 ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu 406

Qy   848 ATGGTGGCAGGTGTTTATCCCGCCAGGCAGTGGGCGCCATCATCCGGCAAGAAGGGCAG 907
          |||||||||||||||||||| |||||||||||||||||||||||||
Db   407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426

Qy   908 CACATCAAACAGCTCTCCGGTTGCCAGCGGCGCCTCCATCAAGATTGACCACCCGAAACA 967
          |||||||||||||||||||:::|||||||||||||||||||||
Db   427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446

Qy   968 CCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACGGCCCAATTCAAGGCT 1027
          |||:::::||| |||||||||||||||||||||:::
Db   447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466

Qy   1028 CAGGGAAGAAATCATGGCAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG 1087
          ||||||||||:::||||||||||||||||||||||||||||||||
Db   467 GlnGlyArgIleTyrGlyLysLysLysGluGluAsnProLysSerProLysGluLeuVal 486

Qy   1088 AAGCTGGAGACCCACCATACGTGTGCCAGCATCAGCAGTGGCGCGGGTCATTGGCAAAGGT 1147
          ||||||||||||| ::: :::
Db   487 LysLeuGluAlaHisHisArgValProSerPheAlaAlaGlyArgValIleGlyLysLysGly 506

Qy   1148 GGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTGGTAGTACCAAGAGAC 1207
          |||||||||:::|||||||||||||||||:::|||||||||||
Db   507 GlyLysThrValAsnGluLeuGlnAsnLeuSerAlaGluValValProArgAspAsp 526

Qy   1208 CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCATGCCAGT 1267
          |||||||:::||||||||||||||||||||:::||||||||||
Db   527 GlnThrProAspGluAsnAspGlnValIleValLysIleIleGlyHisPheTyrAlaCys 546

Qy   1268 CAGATGGCTCAACGGAGAATCCGAGAACATCCTGGCCCAGGTTAAGCAG---CAGCATCAG 1324
          |||:::||||||||| :::||||||||||||||::: ||| |||
Db   547 GlnValAlaGlnArgArgIleArgGluHisLeuLeuThrGlnValLysGlnHisGlnGln 566

Qy   1325 AAGGGA---CAGAGTAACCAGGCCACGGAGGAAG 1360
          ||||| |||||||||||||||||:||||||||
Db   567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 8
US-10-007-700-348
; Sequence 348, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Li, Samual X.
APPLICANT: Kalos, Michael D.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.

; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-348

Alignment Scores:
Pred. No.:   1.26e-121   Length: 579
Score: 1639.00   Matches: 336
Percent Similarity: 83.89%   Conservative: 44
Best Local Similarity: 74.17%   Mismatches: 57
Query Match: 52.70%   Indels: 16
DB: 12   Gaps: 9

US-09-270-437D-5 (1-1708) x US-10-007-700-348 (1-579)

```
Qy    32  CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC   91
              ||||||  :::||   |||||| |||||||||   || ||||| |||||||||||||
Db   133  ARGGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla  152

Qy    92  TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCCGAGGG------  142
              ||||||||||||| |||  |||   ||||||||| ||||||||||||
Db   153  TyrIleProGluThrSerAlaAlaGlnGlnAsnProLeuGlnLysProArgGlyArgArg  172

Qy   143  GGCTTGGCTCTCGGGGTCACCGTGTGGCAGGCCCCAGCC   202
              ||||||||||||||  :::||||||||||||||||||||
Db   173  GlyLeuGlyGlnArgGlySerArgGlnGlySerPro------GlySerValSer  189

Qy   203  AAGCAGCAGCAAGTGGACATCCCCTTCGGCTTCTCGTGTCCCACCAGTATGTGGGTGCC  262
              ||||:::  ||||||||||||||||||||  |||||||||||||||||||||||
Db   190  LysGlnLysProCysAspLeuProLeuAlaThrIleArgAsnIleThrLysGlnLysLys  209

Qy   263  ATTATTGGCAAGGAGGGGGGCCACCATCCGACAACATCACAAAAACAGACCCAGTCCAAGATA  322
              ||||:||| |||||||| ||||||||| |||||||||||||||||::||||| ||||
Db   210  IleIleGlyLysGluGlyGlyHisAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle  229

Qy   323  GACGTGCATAGGAAGGAAGAAGGAACCCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC  382
              |||||||||||||||||||||||||||||||||||||||||||||:::::||||| ||||
Db   230  AspValHisArgLysGluAsnAlaGluGluGluAlaGlnLysSerIleSerLeuSerThr  249

Qy   383  CCTGAGGGCTGCTCCTCCGCTTGTAAGATCTTGGAGATTATGCATAAAGAGGCTAAG   442
              ||||||||||:::||||| ||||||||||||||||||||||||||||:::
Db   250  ProGluGlyThrSerAlaCysLysSerLeuLeuGluIleMetHisLysGluAlaGln  269

Qy   443  GACACCAAAAACGGCTGACGAGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGG  502
              |||| |||||||||||  :::|||||||||||||  |:::::  |||| |||
Db   270  AspIleLysPheThrGluThrLysArgArgAsnLeuLeuLysIleLeuAlaHisAsnAsnPheValGly  289

Qy   503  CGTCTCATTGGCAAGGAGGAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAA  562
              |||||||||||||||||:::||||||||||||||||||||||:::|||||||
Db   290  ArgLeuIleGlyLysGluGluAspGlyLysArgAsnLeuLysIleLysLeuGlyGluGlnAspThrGluThrLys  309

Qy   563  ATCACCAATCTCTCCGTTGCAGACCCTTACAACCCTTTACAACCCTGAGGGACCATCCACTGTG  622
              ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   310  IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal  329
```

```
Qy   623  AAGGGGGCCATCATCGAGAATTGTTGAGGGGCCGAGCAGGAGAAATAATGAAGAAAGTTCGGGAG  682
              ||||| ||||||||||||||||||||:::|||   ||||||||:::|||||
Db   330  LysGlyAsnValGluThrCysAlaAlaLysAlaGluAlaGluIleIleLysLysIleArgGlu  349

Qy   683  GCCTATGAGAATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCTGGCCTGAAC  736
              :::||||||||||||||||||||||||||| :::   |||||||||:::|||||||
Db   350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn  369

Qy   737  CTGGCTGCTGTAGGTCTTTTCCCAGTTCATCAGGCAGCAGTCCCGCG-------CCT  787
              ||| |||||||  :::|||   ||||||
Db   370  LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyPro  389

Qy   788  CCCAGCAGCGTTACTGGGGCTGCTCCTATAGCTCCTTATGCAGGCTCCGAGCAGGAG  847
              ||||| |||||||||||||||||||||:::       :::|||||||
Db   390  ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu  406

Qy   848  ATGGTGCAGGTGTTATCCCGCCAGGACAGTGGGCGCCATCATCGGCAAGAAGGGGCAG  907
              |||||||||||||||||||||||| :::||||| ||||| |||||||
Db   407  ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlyGlyGln  426

Qy   908  CACATCAAACAGCTCTCCGGTTGCCAGCGCGCCTCCATCAAGATTGACCACCCGAAACA  967
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427  HisIleLysGlnLeuSerArgPheAlaGlnAlaSerIleLysIleAspHisProAlaGluAla  446

Qy   968  CCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCGCCAGGAGGCCAATTCAAGGCT  1027
              |||||||||||||||||||||| ||||||||||||||||||:::|||||||||
Db   447  ProAspAlaLysValArgMetValIleIleThrGlyProGlnAlaGlnPheLysAla  466

Qy  1028  CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTGGTCCCAAGGAGGAAGTG  1087
              |||||||| ||||||||||||||||||||||||||||||||:::|||||
Db   467  GlnGlyArgIleTyrGlyLysLeuLysGluGluAsnPheLeuValSerProLysGluGluVal  486

Qy  1088  AAGGTGGAGACCCACATAGTGTGCCAGCATCAGCAGTGGCCGGGTCATTGCAAAGT  1147
              ||||| |||||||| |||||||||||||:::
Db   487  LysLeuGluAlaHisIleArgValProSerPheAlaLysAlaGlyArgValIleGlyLysGly  506

Qy  1148  GGAAAAACGGTGAACGAGTTGCAGAATTTGACGCGAGCTGAGGTGGTAGTACCAAGAGAC  1207
              ||||||||||||||||||||||||||||||||| |||:::|||||||||||||||||||
Db   507  GlyLysThrValAsnGluValGlnAsnLeuSerSerAlaGluValValProArgArgAsp  526

Qy  1208  CAGACCCCTGATGAGAACGACCAGTCATCGTGAAAATCATCGGACATTTCTATGCCAGT  1267
              ||||||||||||||||||||||||||  :::||||| ||:::||||| ||||||
Db   527  GlnThrProAspGluAsnAspGlnSerValValLysIleIleThrGlyThrHisPheTyrAlaCys  546

Qy  1268  CAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAG---CAGCATCAG  1324
              ||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db   547  GlnValAlaGlnArgLysIleGlnIleLeuLeuThrGlnValLysGlnHisGlnGlnGln  566

Qy  1325  AAGGGA---CAGAGTAACCAGGCCCAGGCCACGAGGAAG  1360
              |||||   |||:::||||||||||||||
Db   567  LysAlaLeuGlnSerGlyProProGlnSerArgArgLys  579
```

RESULT 9
US-10-007-700-446
; Sequence 446, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samual X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng

Alignment Scores:
Pred. No.:   579
Score: 1639.00
Percent Similarity: 83.89%   Length: 579
Best Local Similarity: 74.17%   Matches: 336
Query Match: 52.70%   Conservative: 44
DB: 12   Mismatches: 57
                                        Indels: 16
                                        Gaps: 9

US-09-270-437D-5 (1-1708) x US-10-007-700-446 (1-579)

Qy   32 CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGAGAGACCATCCCTGAGAAGGTCTCC   91
Db 133 ArgGlnAlaLeuLeuAspLysLeuAsnGlyHisGlnLeuGluAsnGluThrLeuLysValAla 152

Qy   92 TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCCGAGGG------ 142
Db 153 TyrIleProAspGluThrAlaAlaGlnIleGlnAsnProLeuGlnProLeuGlyProArgGlyArgArg 172

Qy 143 GGGCTTTGGCTCTCGGGGTCAGCCCGCCAGGGGCTCACCTGTGGCAGCGGGGGCCCCAGCC 202
Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnSerPro-------GlySerValSer 189

Qy 203 AAGCAGCAGCAAGTGACATCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCC 262
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209

Qy 263 ATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCCGCAAAACAGACCCAGTCCAAGATA 322
Db 210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229

Qy 323 GACGTGCGATAGGAGGAAGGAACGCAGTGCAGCTGAAAAAAGCCATCAGTGTGCACTCCACC 382
Db 230 AspValHisArgArgLysGluArgThrGlnCysSerLeuLysLysAlaIleSerValLeuSerThr 249

Qy 383 CCTGAGGGCTGCTCCTCCGCTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG 442
Db 250 ProGluGlyCysSerSerAlaAlaCysLysSerSerLeuLeuGluIleMetHisLysGluAlaGln 269

Qy 443 GACACCAAAACGGCTGACGAGGTTCCCTGAGATCCTGGCCCATATAATAACTTTGTAGGG 502
Db 270 AspIleLysPheThrGlyGluIleProGluIleLeuLeuAlaHisAlaAsnPheValGly 289

Qy 503 CGTCTCATTGGCAAGGAGGGACGGAACCTGAGGAAGGTAGAGCGAGCAAGACAAAA 562
Db 290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGluGlnThrGlyThrLys 309

Qy 563 ATCACCAATCTCCGTGTCCAAGACCTTACCCTTACACCCTGAGGAGGACCATCACTGTG 622
Db 310 IleThrIleSerProLeuGlnLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

Qy 623 AAGGGGGCCATCGAGAATTGTTGCAGGGGCCGAGCGGGAAATAATGAAGAAAGTTCGGGAG 682
Db 330 LysGlyAlaIleGluAsnCysCysArgGlyArgGluArgGlyGlnGluGluMetLysLysValArgGlu 349

Qy 683 GCCTATGAGAATGATGTGCTGCCATGAGC-----TCTCACCTGATCCCTGGCCTGAAC 736
Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

Qy 737 CTGGCTGCTGTGTAGGTCTTTTCCCAGCTTCATCCAGCGACGTCCCGCG--------CCT 787
Db 370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro 389

Qy 788 CCCAGCAGCGGTTACTGGGGCTGCTCCTATAGCTCCTTATGCAGGCTCCCGAGCAGGAG 847
Db 390 ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu 406

Qy 848 ATGGTGCAGGTGTTATCCCGCCCAGGCAGTGGGGCCATCATCGGCAAGAAGGGGCAG 907
Db 407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysLysGlyGln 426

Qy 908 CACATCAAACAGCTCCTCCCGGTTGCCAGCGGCCTCATCAAGATTGACAAGCCACCCGAAACA 967
Db 427 HisIleLysGlnLeuLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla 446

Qy 968 CCTGACTCCCAAAGTTCGTATGGTTATCATCGAACTCAGGAGACCGCCAGAGCCCAATTCAAGGCT 1027
Db 447 ProAspAlaLysValArgMetValIleIleProThrProGluAlaGlnPheLysAla 466

Qy 1028 CAGGGAAGAATCTATGGCAACTCAAGGAGGAGAACTTCTTGGTCCCAAGGAGAAGTG 1087
Db 467 GlnGlyArgIleTyrGlyLysIleGlnGluGluAsnPheValSerProLysGluGluVal 486

Qy 1088 AAGCTGGAGACCCACATACGTGCCAGCATCAGCAGCTGGCGGGTCATTGGCAAAGGT 1147
Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506

Qy 1148 GGAAAAACGGTGAACCAGTTGCAGAATTGACGGGACGAGTCGAGGTGTGTAGTACCAAGAGAC 1207
Db 507 GlyLysThrValAsnGlnLeuGlnAsnLeuSerSerAlaGluValAlaValProAlaArgAsp 526

Qy 1208 CAGACCCCTGATGAGAAGACGACCAGGTCATCGTGAAAAATCATCGGACATTCTATGCCAGT 1267
Db 527 GlnThrProAspGluAsnAspGlnValValLysIlePheTyrAlaCys 546

Qy 1268 CAGATGGCTCAACGGAAGATCCGAGACATCCTGCCCAGTTAAGCAG---CAGCATCAG 1324
Db 547 GlnValAlaAlaGlyLysIleGluGlnIleLeuThrGlnValLysGlnHisGlnGln 566

Qy 1325 AAGGGA---CAGAGTAACCAGGCCAGGCACGGAGGAAG 1360
Db 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgArgLys 579

RESULT 10
US-10-007-700-449

Wed Jul 14 08:56:05 2004

us-09-270-437d-5_1.rapb

```
;  LENGTH: 579
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-10-007-700-449

Alignment Scores:
Pred. NO.:              579
Score:             1639.00
Percent Similarity:   83.89%
Best Local Similarity: 74.17%
Query Match:          52.70%
DB:                       12

US-09-270-437D-5  (1-1708) x US-10-007-700-449  (1-579)

Qy    32  CGGGGAGCCATCATGAGAGCTGAATGGCCACCAGTTGAGAACCATGCCCTGAAGGTCTCC   91
              |||||    |||||::::  ||||||||||||||||||||||||||||
Db   133  ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla  152

Qy    92  TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGTCCGAGGG------  142
              ||||||||||      |||  |||      |||||
Db   153  TyrIleProAspGluThrAlaAlaGlnIleAsnProLeuGlnIleAsnProArgGlyArg  172

Qy   143  GGCTTTGGCTCTCGGGTCAGCCCCCAGGGCTACCTGTGGCAGCGGGGGCCCCAGCC    202
              ||||||||             |||      :::     |||||
Db   173  GlyLeuGlyGlnArgGlySerSerArgGluGlySerPro--------GlySerValSer  189

Qy   203  AAGCAGCAAGTGGACATCCCCTTCGGCTTCCTGGTGCCCACCAGTATGTGGGTGCC     262
              |||||::      |||||||:::||||
Db   190  LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla  209

Qy   263  ATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATA  322
              |||      |||||||||||
Db   210  IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle  229

Qy   323  GACGTGCATAGGAAGGAAGGAGAAGCGCAGTGCACCTGAAAAAGCCATCAGTGTGCACTCCACC  382
              ||||||        |:::      ||||
Db   230  AspValHisArgLysGluAlaGlnValAlaAlaGlyIleLysSerIleThrLeuSerThr  249

Qy   383  CCTGAGGGCTGCTCCTCCGGCTGTGAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG  442
              |||||          |||||||||
Db   250  ProGluGlyLysThrSerAlaAlaCysLysSerGluLeuGluMetHisLysGluAlaGln  269

Qy   443  GACACCAAAAACGGCTGACGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTGTAGGG  502
                   |||        :::||||||||||
Db   270  AspIleLysPheThrArgGluLeuIleProLeuLysIleLeuAlaHisAsnAsnPheValGly  289

Qy   503  CGTCTCATTGGCAAGGAAGGAGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAA  562
              :::||||||||||
Db   290  ArgLeuIleGlyLysGluGlyLysGluGlyArgAsnLeuLysLysIleGluGlnAspThrLys  309

Qy   563  ATCACCATCTCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGGAGGACCACTGTG  622
              |||:::      |||||::::||||||||
Db   310  IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal  329

Qy   623  AAGGGGGCCATCGAGAATTGTTGCAGGGCCAGCAGGAAATAATGAAGAAGTTCGGGAG  682
              |||||||||:::|||||||
Db   330  LysGlyAlaAsnValGluThrCysAlaAlaLysAlaGluGluIleIleMetLysLysIleArgGlu  349

Qy   683  GCCTATGAGAATGATGTGGCCTGCCATGAGC------TCTCACCTGATCCCTGGCCTGAAC  736
              ::||||||||::|:::|||:::
Db   350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuLeuIleProGlyLeuAsn  369

Qy   737  CTGGGCTGCTGTGAGGTCTTTTCCCAGCTTCATCCAGCCGAGTCCCCGCCG-------CCT  787
              |||      :::|||      |:::
Db   370  LeuAsnAlaLeuGlyLeuPheLeuGlnThrSerGlyMetProProThrSerGlyPro  389

Qy   788  CCCAGCAGCGTTACTGGGCGTCCTCCTTATGCAGGCTCCCGCAGGAGGAG  847
                   |||||||||||||||
Db   390  ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGlu  406

Qy   848  ATGGTGCAGGTGTTTATCCCCGCCAGGCAGTGGGCGCCCATCATCGCAAGAAGGGGCAG  907
              |||      ::::|||
```

```
Db   407  ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln  426

Qy   908  CACATCAAACAGCTCTCCCGGTTGGCAGCGGCCCTCCATCAAGATTGCACCACCGAAAACA  967
              ||||||||||          |||||||||||||||
Db   427  HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysSerIleAlaProAlaGluAla  446

Qy   968  CCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCCGACCAGGGCCAATTCAAGGCT  1027
              |||||:::||||||||
Db   447  ProAspAlaLysValArgMetValIleIleThrAspProThrProGlnAlaGlnPheLysAla  466

Qy  1028  CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG  1087
              |||||||||||||              |||||||||
Db   467  GlnGlyArgIleTyrGlyLysLeuLysGluGluAsnPheValSerProLysGluGluVal  486

Qy  1088  AAGTCTGAGACCACCATACGTGTGCAACATCAGCAGCTGGCCGGTCATTGGCAAGGT  1147
              |||||          :::
Db   487  LysLeuGluLysHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly  506

Qy  1148  GGAAAAACGGTGACGAGTTGCAGAATTGACGGCCAGCTGAGGTGGTAGTACCAAGAGAC  1207
              ||||||||||||
Db   507  GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgArgAsp  526

Qy  1208  CAGACCCCTGATGAGAGCAGCAGCACCAGGTCATCGTGAAAATCATCGACATTCTATGCCAGT  1267
              |||||||::::      |||||
Db   527  GlnThrProGluGluGlnSerAlaAspGlnValValAlaValLysSerIleThrGlyHisPheTyrAlaCys  546

Qy  1268  CAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGTTAAGCAG---CAGCATCAG  1324
              |||||              |:::
Db   547  GlnValAlaGlnAlaArgLysIleGlnGluIleLeuThrGlnLeuSerGlnValHisGlnGln  566

Qy  1325  AAGGGA---CAGAGTAACCAGCCCCAGGCACGGAGGAAG  1360
              |||||||||
Db   567  LysAlaLeuGlnSerGlyProProGlnSerArgArgLys  579
```

RESULT 11
US-10-117-982-348
; Sequence 348, Application US/10117982
; Publication No. US20030138838A1
; GENERAL INFORMATION:
;   APPLICANT: Foy, Teresa M.
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Vedvick, Thomas S.
;   APPLICANT: Carter, Darrick
;   APPLICANT: Watanabe, Yoshihiro
;   APPLICANT: Henderson, Robert A.
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Mericle, Barbara
;   APPLICANT: Spies, Gregory A.
;   APPLICANT: Fan, Liqun
;   APPLICANT: Wang, Tongtong
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C18
;   CURRENT APPLICATION NUMBER: US/10/117,982
;   CURRENT FILING DATE: 2002-04-05
;   NUMBER OF SEQ ID NOS: 484
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 348
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-117-982-348

Alignment Scores:
Pred. No.:   579
Score: 1639.00
Percent Similarity: 83.89%
Best Local Similarity: 74.17%
Query Match: 52.70%
DB: 14

US-09-270-437D-5 (1-1708) x US-10-117-982-348 (1-579)

RESULT 12
US-10-117-982-446
; Sequence 446, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 446
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-446

Alignment Scores:
Pred. No.:
Score: 1.26e-121   Length: 579
                            1639.00   Matches: 336
Percent Similarity: 83.89%   Conservative: 44
Best Local Similarity: 74.17%   Mismatches: 57
Query Match: 52.70%   Indels: 16
DB: 14   Gaps: 9

US-09-270-437D-5 (1-1708) x US-10-117-982-446 (1-579)

Qy   1325 AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGGAAG 1360
           |||||| |||||||||||||||:|||||||||||
Db   567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 13
US-10-117-982-449
; Sequence 449, Application US/10117982
; Publication No. US20030138418A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-449

Alignment Scores:
Pred. No.:   Length: 579
Score: 1.26e-121   Matches: 336
Percent Similarity: 1639.00   Conservative: 44
Best Local Similarity: 83.89%   Mismatches: 57
Query Match: 74.17%   Indels: 16
DB: 52.70%   Gaps: 9
                         14

US-09-270-437D-5 (1-1708) x US-10-117-982-449 (1-579)

Qy   32 CGGGGAGCCATCATGAAGCTGATGGCCACCAGTTGAGAACCATGCCTGAAGGTTCC 91
           |||||||:|| |||||||||:||||||||||||||:|||||||||:|||||
Db   133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152

Qy   92 TACATCCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCCGAGG------ 142
           |||| |||:|||||||| ||||||||||||||||||||||||
Db   153 TyrIleProAspGluThrAlaGlnGlnAsnGlnGlnProLeuGlnAsnProArgGlyArgArg 172

Qy   143 GGCTTGGCTCTCGGGGTCAGCCCCATGATGGGCTCACCTGTGGGCAGCGGGGGCCCCAGGC 202
           |||| ||||||| |||:::
Db   173 GlyLeuGlyGlnArgGlySerSerSerArgArgGlnGlySerPro-------GlySerValSer 189

Qy   203 AAGCAGCAGCAAGTGGACATCCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCC 262
           ::||||||||||||| |||||||||:
Db   190 LysGlnLysProCysAspGluProLeuArgLeuLeuValProThrGlnPheValGlyAla 209

Qy   263 ATTATTGGCAAGGAGGGGGGCCACCATCGCAACAATCACAACAAGACCAGTCCAAGATA 322
           ||||| |||||||||||||||:|||||||||||||||||||||:|
Db   210 IleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleArgGlnThrSerLysIle 229

Qy   323 ATTATTGGACAAGGAGGGGGGCCACCATCGCAACAATCACAACAAGACCAGTCCAAGATA 322
           ...

Qy   383 CCTCAGGGGCTGCTCCTCCGGTTGAAGATGATCTTGAAGATTATGCATAAAGAGGCTAAG 442
           ...
Db   250 ProGluGlyThrSerAlaAlaCysLysSerSerIleLeuGluIleMetHisLysGluAlaGln 269

Qy   443 GACACCAAAACGGCTGACGAGGTTCCCTGAAGATCCTGACGAGTTGTTGTTAGGG 502
           |||| ||||||||||||||||:

APPLICANT: Kalos, Michael D.
APPLICANT: Mericle, Barbara
APPLICANT: Spies, Gregory A.
APPLICANT: Fan, Liqun
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C18
CURRENT APPLICATION NUMBER: US/10/117,982
CURRENT FILING DATE: 2002-04-05
NUMBER OF SEQ ID NOS: 484
SEQ ID NO: 480
SOFTWARE: FastSEQ for Windows Version 4.0
LENGTH: 579
TYPE: PRT
ORGANISM: Homo sapiens
US-10-117-982-480

Alignment Scores:
Pred. No.:   579
Score: 1.26e-121
Percent Similarity: 1639.00
Best Local Similarity: 83.89%
Query Match: 74.17%
DB: 52.70%
                              14

US-09-270-437D-5 (1-1708) x US-10-117-982-480 (1-579)

Qy   32 CGGGGAGCCATCATGAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC 91
          ||| |||||| ||::: ||||||| |||
Db   133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152

Qy   92 TACATCCCCATGAGCAGATAGCA---CAGGGACCTGAGAATGGCGCGCCAGGG------ 142
          ||| ||||||| ||||| |||||| ||||
Db   153 TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnProArgGlyArgArg 172

Qy   143 GGCTTTGGCTCTCGGGTCGGTCAGCCCGCCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCC 202
          ||:::   :::
Db   173 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro-------GlySerValSer 189

Qy   203 AAGCAGCAGCAAGTGGACATCCCCTTCGGCTCTCGGTGCCACCCAGTATGTGGTGCC 262
          ||::: |:::|||||| ||| |||:::||||
Db   190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla 209

Qy   263 ATTATTGGCAAGGAGGGGCCACCATCCGCAACATCCAAAACAGACCCAGTCCAAGATA 322
          ||| ||| |||:::||| |||
Db   210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleGlnThrGlnSerLysIle 229

Qy   323 GACGTGCATAGGAAGGAAGAGAACGCAGGTGCACGTGAGAAAAGCCATCAGTGTCACTCCACC 382
          ||| |:::|||
Db   230 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThr 249

Qy   383 CCTGAGGGCTGCTCCTCCGGTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG 442
          ||| |||||| |||:::|||
Db   250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaLys 269

Qy   443 GACACCAAAACCGCTGACGAGGTTCCCCTGAAGATCCTGAGAGTGAGAAGGTAGAGACAAAA 502
          ||| |:::|||
Db   270 AspIleLysPheThrGluGluIleProLeuLysIleLeuArgValArgArgLeuGluThr 289

Qy   503 CGTCTCATTGGCAAGGAGGAAGACGGAACCTGAGAAGCTGAAGAAGCTGAAGACCCTGAGAGGA 562
          ||| |:::|||
Db   290 ArgLeuIleGlyLysGluGluAspGlyThrLysLeuLysLysLeuLysThrLeuArgGly 309

Qy   563 ATCACCATCTCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTG 622
          ||| |:::|||
Db   310 IleThrIleSerProLeuGlnThrLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

Qy   623 AAGGGGGCCATCGAGAATTGTTGCAGGGCCAGCAGAAATAATGAAGAAAGTTCGGGAG 682
          |||::: |:::|||
Db   330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu 349

Qy   683 GCCTATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCCTGAAC 736

```
Percent Similarity:      83.89%                    Conservative:   44
Best Local Similarity:   74.17%                    Mismatches:     57
Query Match:             52.70%                    Indels:         16
DB:                      15                        Gaps:            9

US-09-270-437D-5 (1-1708) x US-10-313-986-348 (1-579)
```

```
Qy    32  CGGGGAGCCATCATGAAGTGAATGGCCACCAGTTGGAGAACCATGCCTGAGAGGTCC   91
          ||  ||  ||  ||||:::||||||||||   |||  ||:|||||||||||||||||:::
Db   133  ArgGlnAlaLeuAspIleAspLysLeuAsnGlyPheGlnLeuGlnPheThrLeuLysValAla  152

Qy    92  TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCGAGG------  142
          ||||||||||| ||  ||  |||   ||| |||| ||||| |||| |||
Db   153  TyrIleProAspGluThrAlaAlaGlnAlaGlnProLeuGlnGlnProLeuArgGlyArgArg  172

Qy   143  GGCTTTGGCTCTGGGGTCAGGGCTCCACCTGTGGCAGGGGGGCCCCAGCC          202
          ||||:::||||||||  ||||||  ||:::      |||:::
Db   173  GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro------GlySerValSer  189

Qy   203  AAGCAGCAAGTGGACATCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCC  262
          |||||  :::|||||| || :::|||| ||||||  ||  :::|||  ||||  |||
Db   190  LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnProHisValGlyAla  209

Qy   263  ATTATTGGCAAGGAGGGGGCCACCATCCGCAACACAACAAGACCCAGTCCAAGATA  322
          |||||||| |||||||||||| |||||||||||||||:::|||  ||||||| |||
Db   210  IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle  229

Qy   323  GACGTGCATAGGAGGGAGAACCGAGGTGCAGTGAAAAAGCCATCAGTGTGACTCCACC  382
          |||||| |||||||||| |||||||||||||:::|||| ||||||| ||||:::
Db   230  AspValHisArgArgGluAsnArgGlyAlaValLysLysProSerValThrLeuSerThr  249

Qy   383  CCTGGAGGGTGCTCCTCCGCTGTGAAGATGATCTGAAGATGATTATGCATAAAAGGCTAAG  442
          ||||||||| ||||||||  ||||||  ||||| ||  |||||| |||||||||| |||
Db   250  ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleLeuMetHisLysGluAlaGln  269

Qy   443  GACACCAAAACGGCTGACGAGTTCCCCTGAAGATCCTGGCCCATAATAACTTGTAGG  502
          |||| |||| |||||||||||||  ||:::||||||||||||  ||  ||:::
Db   270  AspIleLysPheLysGluThrGluLeuProLeuLysIleLeuAlaHisAsnAsnPheValGly  289

Qy   503  CGTCTCATTGGCAAGGAGGAGGACCGAACTGAGAAGGTAGAGCAAGATACCGAGACAAAA  562
          ||||||||||||  |||||||||||:::||||||||  ||||| |||||||||:::
Db   290  ArgLeuIleGlyLysGluGluAspArgThrGluLysValGluGlnAspThrGluThrLys  309

Qy   563  ATCACCACTCCTCGTTGCCAAGCACCTTACCCTTACACCCTGAGAGCCATCACTGTG  622
          |||||  :::|||| || ||  |||||| |||||  |||||||||||||||| |||
Db   310  IleThrLeuSerProLeuGlnLeuTyrArgLeuTyrAsnProGluArgThrIleThrVal  329
```

```
Qy   623  AAGGGGGCCATCAGAATTGTTGCAGAGGCCGAGCGAGCAGAAAAATAATGAAGAAAGTTCGGGAG  682
          ||||     :::|||  ||||||||||||||||:::
Db   330  LysGlyAsnValGluThrCysAlaLysAlaGluGluLeuMetLysLysIleArgGlu  349

Qy   683  GCCTATGAGAATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCTGGCCTGAAC  736
          ||||||||||||||||||:::          :::
Db   350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleLeuProGlyLeuAsn  369

Qy   737  CTGCTCGTCTAGGTCTTTTCCCAGCTTCATCCAGCGGCAGTCCCG--------CCT  787
          |||:::|||||||||||||||||||:::
Db   370  LeuAsnAlaLeuGlyLeuPheLeuProProThrGlyMetProProProSerGlyPro  389

Qy   788  CCCAGCAGCGTTACTGGGGCTGTCCTTATGCAGGCTCCCTTATGCAGGCTGTCCCAGGCAGGAG  847
          |||||||||:::            |||||||||||:::
Db   390  ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGln  406

Qy   848  ATGGTGCAGGTGTTTATCCCCGCCACCGCCAAGAGGGGCAG  907
          ||||||||||||||:::
Db   407  ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleGlyAlaIleGlyLysGlnGln  426

Qy   908  CACATCAAACAGCTCTCCCGGTTTGCCAGCGGCCTCCATCAAGATTGCACCACCCGAAACA  967
          ||||:::|||||||||||||||:::
Db   427  HisIleLysGlnLeuSerArgPheAlaLysGlyLysLysProAlaAlaSerIleLeuIleGluAla  446

Qy   968  CCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACGCCCAGAGGCCCAATTCAAGGCT  1027
          |||||||||||||||:::
Db   447  ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysSerAla  466
```

```
Db   350  :::||||||||::::||||||       :::||||||||||||||||||||
         SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleLeuProGlyLeuAsn  369

Qy   737  CTGCTCGTCTAGGTCTTTTCCCAGCTTCATCCAGCGGCCG--------CCT  787
         ||||   ||  :::|||
Db   390  ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu  406

Qy   848  ATGGTGCAGGTGTTTATCCCCGCCACCGCCAAGAGGGGCAG  907
         |||||||||||||||:::
Db   407  ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleGlyAlaIleGlyLysGlnGln  426

Qy   908  CACATCAAACAGCTCTCCCGGTTTGCCAGCGGCCTCCATCAAGATTGCACCACCCGAAACA  967
         ||||:::||||||||||||||||:::
Db   427  HisIleLysGlnLeuSerArgPheAlaLysGlyLysLysProAlaAlaSerIleLeuIleGluAla  446

Qy   968  CCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACGCCCAGAGGCCCAATTCAAGGCT  1027
         |||||||||||||||:::
Db   447  ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysSerAla  466

Qy  1028  CAGGGAGGAATTCATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCAAGGAGGAAGTG  1087
         ||||||||:::
Db   467  GlnGlyArgIleTyrGlyLysLysGluLeuLysGluGluAsnPheValSerProLysGluGluVal  486

Qy  1088  AAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCGCCGGGTCATTGGCAAAGGT  1147
         ||||||:::
Db   487  LysLeuGluLysAlaHisIleArgValProSerPheAlaGlyArgValIleIleGlyLysGly  506

Qy  1148  GGAAAAACGGTGAACGAGTTGACGGCAGAATTGAGCAGCTCGAGGTGGTGTAGCTACCAGAGAC  1207
         ||||||||:::
Db   507  GlyLysThrValAsnGluLeuGlnLeuSerSerAlaGluValValProArgArgAsp  526

Qy  1208  CAGACCCCTGATGAGAACGAACGACCAGTCATCGTGAAAATCATCGGACATTTCTATGCGCAGT  1267
         |||||||:::
Db   527  GlnThrProAspGluAsnGluArgProValIleValLysIleIleGlyHisPheTyrAlaCys  546

Qy  1268  CAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAG---CAGCATCAG  1324
         |||||||:::
Db   547  GlnValAlaGlnArgLysIleGlnAspIleLeuAlaGlnValLysGlnGlnHisGlnGln  566

Qy  1325  AAGGGA--CAGAGTAACCAGGCCCAGGCACGCAGGGAGAAG  1360
         ||||||   ||| |||  :::||||||
Db   567  LysAlaLeuGlnSerGlyProProGlnSerArgArgLys  579
```

RESULT 15
US-10-313-986-348
; Sequence 348, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
;   APPLICANT: Foy, Teresa M.
;   APPLICANT: McNabb, Andria
;   APPLICANT: Watanabe, Yoshihiro
;   APPLICANT: Reed, Steven G.
;   APPLICANT: Wang, Tongtong
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C19
;   CURRENT APPLICATION NUMBER: US/10/313,986
;   CURRENT FILING DATE: 2002-12-04
;   SOFTWARE: FastSEQ for Windows Version 4.0
;   NUMBER OF SEQ ID NOS: 560
; SEQ ID NO 348
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-313-986-348

```
Alignment Scores:
Pred. No.:      1.26e-121      Length:    579
Score:          1639.00        Matches:   336
```

```
Db     447  ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla  466

Qy    1028  CAGGGAAGAATCTATGGCAAACTCAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG  1087
            |||||||||||||||||||:::||||||||||||||||  ||||||||||||||||||
Db     467  GlnGlyArgIleTyrGlyLysLeuArgArgArgThrSerPheValSerGluGluGluVal  486

Qy    1088  AAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGT  1147
            |||||||||||||||||||:::||        ||||||||||||||||||||||||||
Db     487  LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly  506

Qy    1148  GGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTGGTAGTACCAGAGAC  1207
            |||||||||||||||||||::::::||||||||||||||||||||||||||||||||
Db     507  GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluGluValValProArgArgAsp  526

Qy    1208  CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTCTATGCCAGT  1267
            ||||||||||||||||||||::||||||||||||::||  ||||||||||||
Db     527  GlnThrProAspGluAsnAspGlnValValIleValLysIleThrGlyHisPheTyrAlaCys  546

Qy    1268  CAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGTTAAGCAG---CAGCATCAG  1324
            |||::|||||||||||||::::||||||||||||||||||||||||   |||  |||
Db     547  GlnValAlaGlnArgLysIleGlnIleLeuGlnHisHisGlnGln  566

Qy    1325  AAGGGA--CAGAGTAACCAGGCCCAGGCACGGAGGAAG  1360
            ||||||  ||::||||||||
Db     567  LysAlaLeuGlnSerGlyProProGlnSerArgArgLys  579


Search completed: July 13, 2004, 12:27:58
Job time : 110 secs
```

```
OM nucleic - protein search, using frame_plus_n2p model

                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:   July 13, 2004, 12:28:03 ; Search time 22 Seconds
                              (without alignments)
                     8016.105 Million cell updates/sec

Title:          US-09-270-437d-5
Perfect score:  560
Sequence:       1 aggacggctgccgcaccgcc.........atttcctcaggtttaaa 1708

Scoring table:  OLIGO
          Xgapop  60.0 , Xgapext  60.0
          Ygapop  60.0 , Ygapext  60.0
          Fgapop   6.0 , Fgapext   7.0
          Delop    6.0 , Delext    7.0

Searched:       389414 seqs, 5162971 residues
Word size:      1
Total number of hits satisfying chosen parameters:      663654
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=x1p
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_13072004_121956_10024/app_query.fasta_1.1863
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MISMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN_1_33 @runat_13072004_121956_10024 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :   Issued_Patents_AA:*
   1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
   2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
   3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
   4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
   5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
   6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	166	29.6	577	3	US-09-261-855-2	Sequence 2, Appli
2	49	8.8	49	3	US-09-261-855-18	Sequence 18, Appl
3	48	8.6	48	3	US-09-261-855-20	Sequence 20, Appl
3	47	8.4	47	3	US-09-261-855-17	Sequence 17, Appl
4	45	8.0	47	3	US-09-261-855-19	Sequence 19, Appl
5	36	6.4	579	4	US-09-643-597-176	Sequence 176, App
7	36	6.4	579	4	US-09-643-597-348	Sequence 348, App
8	36	6.4	579	4	US-09-480-884A-176	Sequence 176, App
9	36	6.4	579	4	US-09-542-615A-176	Sequence 176, App
10	36	6.4	579	4	US-09-542-615A-348	Sequence 348, App
11	36	6.4	579	4	US-09-606-421B-176	Sequence 176, App
12	36	6.4	579	4	US-09-606-421B-348	Sequence 348, App

```
13  13  28  5.0  47  3  US-09-261-855-21    Sequence 21, Appl
14  14  19  3.4  48  3  US-09-261-855-24    Sequence 24, Appli
15  15  14  2.5  14  3  US-09-261-855-3     Sequence 3, Appli
16  16  14  2.5  14  3  US-09-261-855-46    Sequence 46, Appl
17  17  14  2.5  49  3  US-09-261-855-22    Sequence 22, Appl
18  18  11  2.0  11  3  US-09-261-855-13    Sequence 13, Appl
19  19  11  2.0  47  3  US-09-261-855-23    Sequence 23, Appl
20  20   9  1.6  529 4  US-09-81-656-1      Sequence 1, Appli
21  21   8  1.4  142 4  US-09-252-991A-23673  Sequence 23673, A
22  22   8  1.4  161 4  US-09-252-991A-24938  Sequence 24938, A
23  23   8  1.4  212 4  US-09-252-991A-32417  Sequence 32417, A
24  24   8  1.4  287 4  US-09-252-991A-29951  Sequence 29951, A
25  25   8  1.4  317 4  US-09-252-991A-30984  Sequence 30984, A
26  26   8  1.4  322 4  US-09-252-991A-25412  Sequence 25412, A
27  27   8  1.4  448 4  US-09-543-681A-6550   Sequence 6550, Ap
28  28   8  1.4  2568 4 US-09-866-108A-3      Sequence 3, Appli
29  29   7  1.3  21  6  5368712-7             Patent No. 5368712
30  30   7  1.3  23  4  US-09-093-227-2       Sequence 2, Appli
31  31   7  1.3  30  4  US-09-136-251-7       Sequence 7, Appli
32  32   7  1.3  30  4  US-09-634-496-7       Sequence 7, Appli
33  33   7  1.3  30  4  US-09-635-145A-7      Sequence 7, Appli
34  34   7  1.2  34  4  US-09-079-030-57      Sequence 57, Appl
35  35   7  1.2  35  4  US-08-722-015A-224    Sequence 224, App
36  36   7  1.2  50  3  US-09-261-855-25      Sequence 25, Appl
37  37   7  1.3  55  3  US-09-162-994-14      Sequence 14, Appl
38  38   7  1.3  56  4  US-09-187-789-64      Sequence 64, Appl
39  39   7  1.3  56  4  US-09-139-600-59      Sequence 59, Appl
40  40   7  1.3  57  1  US-08-137-614A-19     Sequence 19, Appl
41  41   7  1.3  63  4  US-08-963-851-18      Sequence 18, Appl
42  42   7  1.2  78  4  US-09-540-236-2668    Sequence 2668, Ap
43  43   7  1.2  85  4  US-09-732-210-1119    Sequence 1119, Ap
44  44   7  1.2  90  4  US-09-040-229B-12     Sequence 12, Appl
45  45   7  1.3  97  4  US-09-621-976-5598    Sequence 5598, Ap
```

ALIGNMENTS

```
RESULT 1
US-09-261-855-2
; Sequence 2, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-2

Alignment Scores:
Pred. No.:                6.4e-156      Length:           577
Score:                    166.00        Matches:          166
Percent Similarity:       100.00%       Conservative:     0
Best Local Similarity:    100.00%       Mismatches:       0
Query Match:              29.64%        Indels:           0
DB:                       3             Gaps:             0

US-09-270-437D-5  (1-1708)  x US-09-261-855-2  (1-577)

Qy  215  GTGGACATCCCCCTTCGGCTCCTGGTGCCACCAGTATGTGGGTGCCATTATTGGCAAG  274
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  194  ValAspIleProLeuArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLys  213

Qy  275  GAGGGGGGCCACCATCCCAACATCACAAAACACCAGTCCAAGATAGAGACGTGCATAGG  334
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214  GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleGluThrCysIleAspValHisArg  233
```

```
Qy  335 AAGGAGGAACGCAGGTCGAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGC 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  234 LysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCys 253

Qy  395 TCCTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACG 454
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  254 SerSerAlaCysLysMetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThr 273

Qy  455 GCTGACGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTGTAGGGCGTCTCATTGGC 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  274 AlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293

Qy  515 AAGGAAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAAATCACCATCTCC 574
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  294 LysGluGlyArgAsnLeuLysLysValGluGlnAspThrGluThrLysIleThrIleSer 313

Qy  575 TCGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGCCATC 634
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  314 SerLeuGlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIle 333

Qy  635 GAGAATTGTTGCAGGGCGCAGGAGAAATAATGAAGAAAGTTCGGGAGGCCTATGAGAAT 694
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  334 GluAsnCysCysArgAlaGluGluLysIleMetLysLysValArgGluAlaTyrGluAsn 353

Qy  695 GATGTGGCTGCCATGAGC 712
          |||||||||||||
Db  354 AspValAlaAlaMetSer 359
```

RESULT 2
US-09-261-855-18
; Sequence 18, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-18

```
Alignment Scores:
Pred. No.:                8.95e-40      Length:         49
Score:                    49.00         Matches:        49
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              8.75%         Indels:         0
DB:                       3             Gaps:           0
```

US-09-270-437D-5 (1-1708) x US-09-261-855-18 (1-49)

```
Qy  476 ATCCTGGCCCATAATAACTTGTAGGGCGTCTCATTGGCAAGGAGGACGGAACCTGAAG 535
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 IleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLys 20

Qy  536 AAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTT 595
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 LysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeuThrLeu 40

Qy  596 TACAACCCTGAGAGGACCATCACTGTG 622
          |||||||||||||||||
Db   41 TyrAsnProGluArgThrIleThrVal 49
```

RESULT 3
US-09-261-855-20
; Sequence 20, Application US/09261855A

; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-20

```
Alignment Scores:
Pred. No.:                8.8e-39       Length:         48
Score:                    48.00         Matches:        48
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              8.57%         Indels:         0
DB:                       3             Gaps:           0
```

US-09-270-437D-5 (1-1708) x US-09-261-855-20 (1-48)

```
Qy 1103 ATACGTGCCAGATCAGCAGTCTGGCCGGGTCATTGGCAAAAGTGGAAAAACGGTGAAC 1162
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 IleArgValProAlaSerAlaAlaGlyArgValIleGlyLyGlyLysThrValAsn 20

Qy 1163 GAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTACCCAAGAGACCCAGACCCCTGATGAG 1222
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 GluLeuGlnAsnLeuThrAlaAlaGluValValValProArgArgProGlnThrProAspGlu 40

Qy 1223 AACGACCAGTCATCGTGAAAATC 1246
          |||||||||||||||||
Db   41 AsnAspGlnValIleValLysIle 48
```

RESULT 4
US-09-261-855-17
; Sequence 17, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-17

```
Alignment Scores:
Pred. No.:                8.65e-38      Length:         47
Score:                    47.00         Matches:        47
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              8.39%         Indels:         0
DB:                       3             Gaps:           0
```

US-09-270-437D-5 (1-1708) x US-09-261-855-17 (1-47)

```
Qy  233 CTCCTGCTGGTGCCCACCCAGTAGTGGGTGCCATTATTGGCAAGGAGGAGGGGCCACCATCCGC 292
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 LeuLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 20

Qy  293 AACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAGGAAGAACGCAGGTGCA 352
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        21 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 40

Qy       353 GCTGAAAAGCCATCAGTGTG 373
             ||||||||||||||||||||
Db        41 AlaGluLysAlaIleSerVal 47

RESULT 5
US-09-261-855-19
; Sequence 19, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
;   APPLICANT: Ross, Jeffrey
;   TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
;   TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
;   FILE REFERENCE: 960296.95131
;   CURRENT APPLICATION NUMBER: US/09/261,855A
;   CURRENT FILING DATE: 1999-03-03
;   NUMBER OF SEQ ID NOS: 46
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 47
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-261-855-19

Alignment Scores:
Pred. No.:              8.33e-36          Length:          47
Score:                  45.00             Matches:         45
Percent Similarity:     100.00%           Conservative:    0
Best Local Similarity:  100.00%           Mismatches:      0
Query Match:            8.04%             Indels:          0
DB:                                       Gaps:            0

US-09-270-437D-5 (1-1708) x US-09-261-855-19 (1-47)

Qy       857 GTGTTTATCCCCGCCCAGGCAGTGGGGCCATCATCCGCAAGAAGGGGCAGCACATCAAA 916
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 ValPheIleProAlaGlnAlaValGlyAlaIleIleArgLysLysGlyGlnHisIleLys 20

Qy       917 CAGCTCTCCCGGTTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCC 976
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        21 GlnLeuSerArgPheAlaSerAlaSerIleLysIleAlaProProGluThrProAspSer 40

Qy       977 AAAGTTCGTATGGTT 991
             |||||||||||||||
Db        41 LysValArgMetVal 45

RESULT 6
US-09-643-597-176
; Sequence 176, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Li, Samuel X.
;   APPLICANT: Wang, Aijun
;   APPLICANT: Skeiky, Yasir A.W.
;   APPLICANT: Henderson, Robert A.
;   APPLICANT: McNeill, Patricia D.
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C11
;   CURRENT APPLICATION NUMBER: US/09/643,597
;   CURRENT FILING DATE: 2000-08-21
;   NUMBER OF SEQ ID NOS: 369
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 176
;   LENGTH: 579
;   TYPE: PRT
```

```
;   ORGANISM: Homo sapiens
US-09-643-597-176

Alignment Scores:
Pred. No.:              5.28e-27          Length:          579
Score:                  36.00             Matches:         36
Percent Similarity:     100.00%           Conservative:    0
Best Local Similarity:  100.00%           Mismatches:      0
Query Match:            6.43%             Indels:          0
DB:                                       Gaps:            0

US-09-270-437D-5 (1-1708) x US-09-643-597-176 (1-579)

Qy       254 GTGGGTGCCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy       314 TCCAAGATAGACGTGCATAGGAAGGAAGAGAACCGCAGGTGCAGCTGAAAAA 361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGlyAlaGluLys 242

RESULT 7
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Li, Samuel X.
;   APPLICANT: Wang, Aijun
;   APPLICANT: Skeiky, Yasir A.W.
;   APPLICANT: Henderson, Robert A.
;   APPLICANT: McNeill, Patricia D.
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C11
;   CURRENT APPLICATION NUMBER: US/09/643,597
;   CURRENT FILING DATE: 2000-08-21
;   NUMBER OF SEQ ID NOS: 369
;   SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-643-597-348

Alignment Scores:
Pred. No.:              5.28e-27          Length:          579
Score:                  36.00             Matches:         36
Percent Similarity:     100.00%           Conservative:    0
Best Local Similarity:  100.00%           Mismatches:      0
Query Match:            6.43%             Indels:          0
DB:                                       Gaps:            0

US-09-270-437D-5 (1-1708) x US-09-643-597-348 (1-579)

Qy       254 GTGGGTGCCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       207 ValGlyAlaIleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy       314 TCCAAGATAGACGTGCATAGGAAGGAAGAGAACCGCAGGTGCAGCTGAAAAA 361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGlyAlaGluLys 242

RESULT 8
US-09-480-884A-176
; Sequence 176, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
```

```
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Hosken, Nancy A.
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Fanger, Gary R.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C6
;   CURRENT APPLICATION NUMBER: US/09/480,884A
;   CURRENT FILING DATE: 2001-08-27
;   NUMBER OF SEQ ID NOS: 330
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   SEQ ID NO 176
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-480-884A-176

Alignment Scores:
Pred. No.:              5.28e-27         Length:              579
Score:                  36.00            Matches:             36
Percent Similarity:     100.00%          Conservative:        0
Best Local Similarity:  100.00%          Mismatches:          0
Query Match:            6.43%            Indels:              0
DB:                     4                Gaps:                0

US-09-270-437D-5 (1-1708) x US-09-480-884A-176 (1-579)

Qy    254 GTGGGTGCCATTATTGGCAAGGAGGGGGGCCACCATCCGCAACATCACAAAAACAGACCCAG 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 ValGlyAlaIleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy    314 TCCAAGATAGACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAAAAA 361
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 9
US-09-542-615A-176
;   Sequence 176, Application US/09542615A
;   Patent No. 6518256
;   GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy A.
;   APPLICANT: Fanger, Gary R.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C8
;   CURRENT APPLICATION NUMBER: US/09/542,615A
;   CURRENT FILING DATE: 2000-04-14
;   NUMBER OF SEQ ID NOS: 350
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   SEQ ID NO 176
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-542-615A-176

Alignment Scores:
Pred. No.:              5.28e-27         Length:              579
Score:                  36.00            Matches:             36
Percent Similarity:     100.00%          Conservative:        0
Best Local Similarity:  100.00%          Mismatches:          0
Query Match:            6.43%            Indels:              0
DB:                     4                Gaps:                0

US-09-270-437D-5 (1-1708) x US-09-542-615A-176 (1-579)

Qy    254 GTGGGTGCCATTATTGGCAAGGAGGGGGGCCACCATCCGCAACATCACAAAAACAGACCCAG 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 ValGlyAlaIleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy    314 TCCAAGATAGACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAAAAA 361
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 10
US-09-542-615A-348
;   Sequence 348, Application US/09542615A
;   Patent No. 6518256
;   GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy A.
;   APPLICANT: Fanger, Gary R.
;   TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C8
;   CURRENT APPLICATION NUMBER: US/09/542,615A
;   CURRENT FILING DATE: 2000-04-14
;   NUMBER OF SEQ ID NOS: 350
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   SEQ ID NO 348
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-542-615A-348

Alignment Scores:
Pred. No.:              5.28e-27         Length:              579
Score:                  36.00            Matches:             36
Percent Similarity:     100.00%          Conservative:        0
Best Local Similarity:  100.00%          Mismatches:          0
Query Match:            6.43%            Indels:              0
DB:                     4                Gaps:                0

US-09-270-437D-5 (1-1708) x US-09-542-615A-348 (1-579)

Qy    254 GTGGGTGCCATTATTGGCAAGGAGGGGGGCCACCATCCGCAACATCACAAAAACAGACCCAG 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 ValGlyAlaIleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy    314 TCCAAGATAGACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAAAAA 361
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242

RESULT 11
US-09-606-421B-176
;   Sequence 176, Application US/09606421B
;   Patent No. 6531315
;   GENERAL INFORMATION:
;   APPLICANT: Wang, Tongtong
;   APPLICANT: Fan, Liqun
;   APPLICANT: Kalos, Michael D.
;   APPLICANT: Bangur, Chaitanya S.
;   APPLICANT: Hosken, Nancy
;   APPLICANT: Fanger, Gary R.
;   APPLICANT: Li, Samuel X.
;   APPLICANT: Wang, Aijun
;   APPLICANT: Skeiky, Yasir A.W.
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C9
;   CURRENT APPLICATION NUMBER: US/09/606,421B
;   CURRENT FILING DATE: 2000-06-28
;   NUMBER OF SEQ ID NOS: 358
;   SOFTWARE: FastSEQ for Windows Version 3.0
;   SEQ ID NO 176
;   LENGTH: 579
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-606-421B-176
```

Alignment Scores:
Pred. No.:
Score: 5.28e-27
                           36.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 6.43%
DB:
Length: 579
Matches: 36
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-606-421B-176 (1-579)

```
Qy  254 GTGGGTGCCATTATTGGCAAGGAGGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  207 ValGlyAlaIleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy  314 TCCAAGATAGACGTGCATAGGAAGGAGGAACGCAGGTGCAGCTGAAAAA 361
        |||||||||||||||||||||||||||||||||||||||||||
Db  227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242
```

RESULT 12
US-09-606-421B-348
; Sequence 348, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-348

Alignment Scores:
Pred. No.:
Score: 5.28e-27
                           36.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 6.43%
DB:
Length: 579
Matches: 36
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-606-421B-348 (1-579)

```
Qy  254 GTGGGTGCCATTATTGGCAAGGAGGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 313
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  207 ValGlyAlaIleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGln 226

Qy  314 TCCAAGATAGACGTGCATAGGAAGGAGGAACGCAGGTGCAGCTGAAAAA 361
        |||||||||||||||||||||||||||||||||||||||||||
Db  227 SerLysIleAspValHisArgLysGluAsnAlaGlyAlaAlaGluLys 242
```

RESULT 13
US-09-261-855-21
; Sequence 21, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-21

Alignment Scores:
Pred. No.:
Score: 6.01e-19
                           28.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 5.00%
DB:
Length: 47
Matches: 28
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-261-855-21 (1-47)

```
Qy  278 GGGGCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAG 337
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16 GlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLys 35

Qy  338 GAGAACGCAGTGCAGCTGAAAAA 361
        |||||||||||||||||||
Db   36 GluAsnAlaGlyAlaAlaGluLys 43
```

RESULT 14
US-09-261-855-24
; Sequence 24, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-24

Alignment Scores:
Pred. No.:
Score: 5.04e-10
                           19.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.39%
DB:
Length: 48
Matches: 19
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-261-855-24 (1-48)

```
Qy 1121 GCAGCTGCCGCGGTCATTGGCAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTTG 1177
         |||||||||||||||||||||||||||||||||||||||||||
Db    7 AlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeu 25
```

RESULT 15
US-09-261-855-3
; Sequence 3, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46

```
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 3
;  LENGTH: 14
;  TYPE: PRT
;  ORGANISM: Mus musculus
US-09-261-855-3

Alignment Scores:
Pred. No.:              5.27e-05         Length:          14
Score:                  14.00            Matches:         14
Percent Similarity:     100.00%          Conservative:    0
Best Local Similarity:  100.00%          Mismatches:      0
Query Match:            2.50%            Indels:          0
DB:                     3                Gaps:            0

US-09-270-437D-5 (1-1708) x US-09-261-855-3  (1-14)

Qy    134 CGCCGAGGGGGCTTTGGCTCTCGGGGTCAGCCCCGCCAGGGC 175
          ||||||||||||||||||||||||||||||||||||||||||
Db      1 ArgArgGlyGlyPheGlySerArgGlyGlnProArgGlnGly 14

Search completed: July 13, 2004, 12:36:38
Job time : 28 secs
```

OM nucleic - protein search, using frame_plus_n2p model

```
                GenCore version 5.1.6
       Copyright (c) 1993 - 2004  Compugen Ltd.
```

Run on: July 13, 2004, 11:51:37 ; Search time 82 Seconds
    (without alignments)
    11770.516 Million cell updates/sec

Title: US-09-270-437D-5
Perfect score: 3110
Sequence: 1 aggacgctgcgcaccgcc.........atttcctcaggtttaaaa 1708

Scoring table: BLOSUM62
  Xgapop 10.0 , Xgapext 0.5
  Ygapop 10.0 , Ygapext 0.5
  Fgapop 6.0 , Fgapext 7.0
  Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
  Maximum Match 100%
  Listing first 45 summaries

Command line parameters:
```
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2/1/USPTO_spool_p/US09270437/runat_13072004_121921_9584/app_query.fasta_1.1863
-DB=A_Geneseq_29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bts -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN_1_1_154 @runat_13072004_121921_9584 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Database :
```
A_Geneseq_29Jan04:*
 1: geneseqp1980s:*
 2: geneseqp1990s:*
 3: geneseqp2000s:*
 4: geneseqp2001s:*
 5: geneseqp2002s:*
 6: geneseqp2003as:*
 7: geneseqp2003bs:*
 8: geneseqp2004s:*
```

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	2208	71.0	577	2	AAY30649	Aay30649 A murine
2	1639	52.7	579	5	AAB11365	Aab11365 Human lun
3	1639	52.7	579	5	ABB75053	Abb75053 Human lun
4	1639	52.7	579	5	ABB74997	Abb74997 Human lun
5	1639	52.7	579	5	ABB75054	Abb75054 Human lun
6	1639	52.7	579	5	ABP61917	Abp61917 Human lun
7	1639	52.7	579	5	ABP61974	Abp61974 Human lun
8	1639	52.7	579	5	ABP61973	Abp61973 Human lun
9	1639	52.7	579	7	ADA28536	Ada28536 Recombina
10	1639	52.7	579	7	ADA28539	Ada28539 Recombina
11	1639	52.7	579	7	ADA28438	Ada28438 Human lun
12	1639	52.7	586	5	ABB75048	Abb75048 Human lun
13	1639	52.7	586	5	ABP61968	Abp61968 Human lun
14	1639	52.7	586	7	ADA28517	Ada28517 Recombina
15	1637	52.6	579	3	AAB11328	Aab11328 Human lun
16	1637	52.6	579	5	ABB74960	Abb74960 Human lun
17	1637	52.6	579	5	ABP61880	Abp61880 Human lun
18	1637	52.6	579	7	ADA28266	Ada28266 Human lun
19	1637	52.6	579	7	ADD14066	Add14066 Human src
20	1637	52.6	579	7	ADE53471	Ade53471 Human lun
21	1585	51.0	319	4	AAM93826	Aam93826 Human pol
22	1572	50.5	619	4	ABG21963	Abg21963 Novel hum
23	1530.5	49.2	614	4	ABG06794	Abg06794 Novel hum
24	1530.5	49.2	620	4	AAU16163	Aau16163 Human nov
25	1530.5	49.2	620	6	ABU55232	Abu55232 Human nov
26	1518	48.8	583	4	ABG12592	Abg12592 Novel hum
27	1464	47.1	556	5	ABG96346	Abg96346 Human ova
28	1464	47.1	594	5	ABG06795	Abg06795 Novel hum
29	1355.5	43.6	555	6	ABU89799	Abu89799 Novel hum
30	1258	40.5	261	4	AAU16161	Aau16161 Human nov
31	1258	40.5	261	6	ABU55230	Abu55230 Human nov
32	1228	39.5	250	4	AAU16579	Aau16579 Human nov
33	1228	39.5	250	6	ABU55648	Abu55648 Human nov
34	858	27.6	558	6	ABB58367	Abb58367 Drosophil
35	653	21.0	171	4	AAU16166	Aau16166 Human nov
36	653	21.0	171	6	ABU55235	Abu55235 Human nov
37	605.5	19.5	187	6	AAO23971	Aao23971 Human IGF
38	510	16.4	192	4	AAU16164	Aau16164 Human nov
39	510	16.4	192	6	ABU55233	Abu55233 Human nov
40	502	16.1	171	6	AAU16583	Aau16583 Human nov
41	502	16.1	171	6	ABU55652	Abu55652 Human nov
42	465.5	15.0	148	4	ABG21962	Abg21962 Novel hum
43	410	13.2	266	4	ABG12593	Abg12593 Novel hum
44	409.5	13.2	209	4	ABG21961	Abg21961 Novel hum
45	292	9.4	93	4	AAM38501	Aam38501 Peptide #

ALIGNMENTS

```
RESULT 1
AAY30649
ID   AAY30649 standard; protein; 577 AA.
AC
XX
DT   17-NOV-1999  (first entry)
XX
DE   A murine c-myc coding region determinant binding protein.
XX
KW   c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;
KW   endonucleolytic attack; half-life; breast cancer; colon cancer;
KW   pancreatic cancer.
XX
OS   Mus musculus.
XX
PN   WO9946594-A2.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999;  99WO-US004897.
XX
PR   09-MAR-1998;  98US-0077372P.
XX
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ross J;
XX
DR   WPI; 1999-551506/46.
DR   N-PSDB; AAZ10617.
XX
PT   Diagnosing presence or absence of a tumor in a human by examining c-myc
PT   coding region determinant-binding protein.
XX
```

```
PS    Example; Fig 1A-D; 79pp; English.
XX
CC    The present sequence represents a murine c-myc coding region determinant
CC    binding protein (CRD-BP). The presence or absence of a tumor can be
CC    determined by determining the levels of CRD-BP present in the suspect
CC    tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack
CC    and so prolongs its half-life. The methods are used for diagnosing
CC    presence or absence of a tumor in a human, especially breast, colon and
CC    pancreatic cancer. They are also used to inhibit cancer cell growth
XX
SQ    Sequence 577 AA;
```

Alignment Scores:

Pred. No.:	9.76e-198	Length:	577
Score:	2208.00	Matches:	438
Percent Similarity:	98.65%	Conservative:	1
Best Local Similarity:	98.43%	Mismatches:	4
Query Match:	71.00%	Indels:	2
DB:		Gaps:	1

US-09-270-437D-5 (1-1708) x AAY30649 (1-577)

```
Qy    32  CGGGGAGCCATCATGAATGGCCACCAGTTGGAGGAACCATGCCCTGAAGGTCTCC   91
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   133  ArgGlnAlaIleIleMetLysLeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSer  152

Qy    92  TACATCCCGATGAGCAGATAGCACAGGGAACCTGAGAATGGGCGCCGAGGGGCTTTGGC  151
Db   153  TyrIleProAspGluGlnIleThrGlnGlyGluProGluAsnGlyArgArgGlyTyrGlyPheGly  172

Qy   152  TCTCGGGGTCAGCCCCCCCAGGGCTTCACCTGTGGCAGGGGCCCCAGCCAGCAGCAG  211
Db   173  SerArgGlyGlnProArgGlnGlyPheSerProValAlaAlaGlyAlaProAlaLysGlyGln  192

Qy   212  CAAGTGACATCCCCTTCGGCTCCTGGTCGTCCACCAGTATGTGGTGCCATTATTGGC  271
Db   193  ProValAspIleProLeuArgLeuLeuValLeuValProThrGlnTyrValGlyAlaIleIleGly  212

Qy   272  AGGAGGGGGCCAACATCCGCAACATCACAAAAACAGACCCAGTCCAAGATAGACGTGCAT  331
Db   213  LysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerGlnSerValHis  232

Qy   332  AGGAAGAGAACGCAGTGCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGC  391
Db   233  ArgLysGluAsnAlaGlyAlaAlaGlyLysAlaIleSerValHisSerThrProGluGly  252

Qy   392  TGCTCCTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGAGGCTAAGGACCACCAAA  451
Db   253  CysSerSerAlaCysLysMetIleLeuGluIleMetHisLysGluArgLeuArgThrThrLys  272

Qy   452  ACGGCTGACGAGGTTCCCCTGAAGATCCTGGCCCCATAAATAACTTTGTAGGGCCGTCTCATT  511
Db   273  ThrAlaAspGluValProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIle  292

Qy   512  GGCAAGGAAGGACGGGAACCTGAAGAAGGTGAGCAAGAATACCGAGACAAAAATCACCATC  571
Db   293  GlyLysGluGlyGluProGluGluGlyValGluValGlnLysArgLysSerThrIleThr  312

Qy   572  TCCTCGTTGCAAGGACCTTACCCTTTACAACCCTGAGAGAAGGAAGTTAGCTTATCGG  631
Db   313  SerSerLeuGlnAspLeuThrLeuTyrAsnProGluIleArgThrIleLysGlyAla  332

Qy   632  ATCGAGAATTGTTGCAGGGCCGAGCGAGAAATAATGAAGAAAAGTTCGGAGGGCCCATGAG  691
Db   333  IleGluAsnCysCysArgAlaGluGlnGlnLeuMetLysLysValArgGluAlaTyrGlu  352

Qy   692  AATGAATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCTGACCCTGGCTGCT  745
Db   353  AsnAspValAlaAlaMetSerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAla  372

Qy   746  GTAGGTCTTTTTCCCAGCTTCATCCAGCGCAGTCCCGGCCGCTCCGGCAGCGTTACTGGG  805
Db   373  ValGlyLeuPheProAlaSerSerAlaValProProProProSerValThrGly  392

Qy   806  GCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGGATGGTCAGGTGTTTATC  865
Db   393  AlaAlaProTyrSerSerPheMetGlnAlaProGluGlnGluMetValGlnValPheIle  412

Qy   866  CCCGCCAGGCAGTGGGCGCCATCATCGGCAAGAGAGGGGCAGCATCAAACAGCTCTCC  925
Db   413  ProLeuGlnAlaValGlyAlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSer  432

Qy   926  CGGTTTGCCAGCGGCCTCATCAGATTGCACCAAGATTCCACCACCACTGACTCCAAAGTTCGT  985
Db   433  ArgPheAlaSerArgAlaSerIleLysSerIleAlaProProGluThrProAspSerLysValArg  452

Qy   986  ATGGTTATCATCACTGGACCGCCAGAGCCCAATTCAAGGCTTCAGGGAGAAATCTATGC 1045
Db   453  MetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrArgGly  472

Qy  1046  AAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGCTGAAGCTCGAGAACCCACATA 1105
Db   473  LysLeuLysGluGluAsnPhePheGlyProLysGluGluAlaGluValLysLeuGluThrHisIle  492

Qy  1106  CGTGTGCCAGCATCAGCAGTCTGGCCGGGTCATTGGCAAAGGTGGAAAAACGGTGAACGAG 1165
Db   493  ArgValProAlaSerArgAlaAlaGlyArgValIleGlyLysGlyLysLysThrValAsnGlu  512

Qy  1166  TTGCAGAATTGACGGCAGCTGAGGTGGTAGTACCAGAGACCAGACCCCTGATGAGAAC 1225
Db   513  LeuGlnAsnLeuThrAlaAlaGluValValValProGluThrArgAspSerGluAsn  532

Qy  1226  GACCAGGTCATCGTGAAAATCATCGGACCATTTCTATCGCCAGTAGGCTCAGATGGCTCAACGGAAG 1285
Db   533  AspGlnValIleValLysIleIleGlyProPheTyrAlaSerGlnMetAlaGlnArgLys  552

Qy  1286  ATCCGAGACCATCCTGGCCCAGGTTAAGCAGCATCAGAGGACCAGCAGTAACCAGGCC 1345
Db   553  IleArgAspIleLeuAlaGlnValLysGlnHisGlnArgThrSerSerAsnLeuAla  572

Qy  1346  CAGGCCACGGAGGAAG 1360
Db   573  GlnAlaAlaArgLys  577
```

```
RESULT 2
AAB11365
ID   AAB11365 standard; protein; 579 AA.
XX
AC   AAB11365;
XX
DT   21-FEB-2001 (first entry)
XX
DE   Human lung cancer associated antigen L523S.
XX
KW   Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW   vaccine; detection.
XX
OS   Homo sapiens.
XX
PN   WO200061612-A2.
XX
PD   19-OCT-2000.
XX
PF   03-APR-2000; 2000WO-US008896.
XX
PR   02-APR-1999;  99US-00285479.
PR   17-DEC-1999;  99US-00466396.
PR   30-DEC-1999;  99US-00476496.
PR   10-JAN-2000;  2000US-00480884.
PR   22-FEB-2000;  2000US-00510376.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Fan L;
XX
DR   WPI; 2000-628399/60.
```

DR   N-PSDB; AAC66035.
XX
PT   Isolated polypeptide comprising an immunogenic portion of a lung tumor
PT   protein is used for detecting and monitoring progression of lung cancer
PT   in a patient.
XX
PS   Claim 3; Page 259-261; 261pp; English.
XX
CC   This invention describes a novel isolated polypeptide (I) which
CC   comprising an immunogenic portion of a lung tumor protein or variant (P2)
CC   which have cytostatic activity. The polypeptides and polynucleotides are
CC   used in compositions and vaccines to inhibit the development of cancer,
CC   especially lung cancer, in a patient. Methods described in the invention
CC   can be used to monitor the progression of a cancer by carrying out the
CC   detection at subsequent time points and comparing the results from the
CC   different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC   are treated with P2, polynucleotides encoding P2 or antigen presenting
CC   cells expressing P2 and then administered to the patient to inhibit
CC   development of cancer
XX
SQ   Sequence 579 AA;

Alignment Scores:
Pred. No.:   2.36e-144   Length: 579
Score: 1639.00   Matches: 336
Percent Similarity: 83.89%   Conservative: 44
Best Local Similarity: 74.17%   Mismatches: 57
Query Match: 52.70%   Indels: 16
DB: 3   Gaps: 9

US-09-270-437D-5 (1-1708) x AAB11365 (1-579)

Qy   32   CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGAGAACCATGCCCTGAAGGTCTCC   91
         |||||| ::: ||||||||||| |||||||
Db   133   ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla   152

Qy   92   TACATCCCGATGAGCAGATAGCA--CAGGGACCTGAGAATGGGCGCCGAGGG------   142
         |||||| ||| ||| |||||
Db   153   TyrIleProAspGluThrArgAlaAlaGlnAsnProLeuGlnProArgGlyArgArg   172

Qy   143   GGCTTTGGCTCTCGGGGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCC   202
         ||||||||||| ||| ||| :::
Db   173   GlyLeuGlyLyArgGlyGlySerSerArgGlnGlySerPro--------GlySerValSer   189

Qy   203   AAGCAGCAGCAAGTGGACATCCCCCTTCGGCTCCTGCCCACCAGTATGTGGGTGCC   262
         |||||: ::: ||||||||||||||||| ::: |||||||||| |||||
Db   190   LysGlnLysProCysAspLeuProLeuArgLeuLeuProThrGlnPheValGlyAla   209

Qy   263   ATTATTGGCAAGAGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATA   322
         ||||||: ||||||||||||||||||||:|
Db   210   IleIleGlyLysGluGluGlyHisHisArgAsnIleThrLysGlnThrGlnSerLysIle   229

Qy   323   GACGTGCATAGGAGAAGGAGAACCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC   382
         |||||||||||||||||||||||||||:::
Db   230   AspValHisArgArgArgThrArgCysSerLeuLysLysProSerValCysThr   249

Qy   383   CCTGAGGGCTGCTCCTCCGCTTGTAAGATGACTCTGAAGATTATGCATAAAGAGGCTAAG   442
         ||||||| |||| ||||||||||||||||||:::
Db   250   ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln   269

Qy   443   GACACCAAAACGCTGACGAGGTTCCCCTGAAGATCCTGCCCATAATAACTTTGTAGGG   502
         ||| :::|||||||||||||||||||||||||||
Db   270   AspIleLysPheThrGluThrGluLeuLluGlIleLeuLysIleLeuAlaHisAsnAsnPheValGly   289

Qy   503   CGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAA   562
         ||| ||||||||||||||||||||||||||
Db   290   ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysValGluGlnAspThrGluThrLys   309

Qy   563   ATCACCATCTCCCGTTGCCAAGACCTTACCCCTTTACACCCGAGAGACCATCACTGTG   622
         ||| |:::||||||||||||||||||
Db   310   IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal   329

Qy   623   AAGGGGGCCATCGAGATTGTTGCAGGGCCGAGGAAAATAATGAAGAAAGTTCGGGAG   682

Db   330 |||||| ::: |||| ::::|||||||||
         LysGlyGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu   349

Qy   683   GCCTATGAGAATGATGGGTGCCATGAGC-----TCTCACCTGATCCCTGGCCTGAAC   736
         :::||||||||||:::||
Db   350   SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn   369

Qy   737   CTGGCTGCTGTAGGGTCTTTTCCAGCTTCATCCAGCGACGAGTCCCCGCCG-------CCT   787
         |||||:::|||||| :::
Db   370   LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro   389

Qy   788   CCCAGCAGCGTTACTGGGGGCTGCTCCCTATAGCTCCTTATGCAGGCTCCGAGCAGGAG   847
         |||||:::|||| ||||| :::
Db   390   ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu   406

Qy   848   ATGGTGCAGGTGTTTATCCCCGCCCAGGCAGTGGGCGGCCATCATCGGCAAGAAGGGGCAG   907
         |||| ::: |||
Db   407   ThrValHisLeuPheLeuProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln   426

Qy   908   CACATCAAACAGCTCTCCCGTTGCAGCGCCCTCCATCAAGATTGCACCACCCGAAACA   967
         |||||||||||||||||||||||:|| |||
Db   427   HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla   446

Qy   968   CCTGACTCCAAGTTCGTATGTTATCATCACTGACCGCCAGAGGCCCAATTCAAGGCT   1027
         |||||||||||||||||||||||||||||||||||
Db   447   ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla   466

Qy   1028   CAGGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTGGTCCAAGGAGAAGTG   1087
         |||||||| |||||||||||||
Db   467   GlnGlyArgIleTyrGlyLysIleLysGluGlnAsnPheValSerProLysProGluVal   486

Qy   1088   AAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTCGGCCGGGTCATTGGCAAAGGT   1147
         |||||| :::
Db   487   LysLeuGluAlaHisIleArgValProSerPheAlaGlyAlaGlyArgValIleGlyLysGly   506

Qy   1148   GGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTGTGAGTACCAAGAGAC   1207
         ||||||||| |||
Db   507   GlyLysThrValAsnGluGlnAsnLeuSerSerAlaGluValValProArgArgAsp   526

Qy   1208   CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGT   1267
         |||||||||||||||||||||||||||||||||
Db   527   GlnThrProAspGluAsnAspGlnValValIleValLysIleIleThrGlyHisPheTyrAlaCys   546

Qy   1268   CAGATGGCTCAACGGAAGAATCCGAGACATCCTGGCCCAGGTTAAGCAG---CAGCATCAG   1324
         |||||||::|||||||||||:::
Db   547   GlnValAlaGlnArgLysIleLeuGlnGluLeuThrGlnValLysGlnHisGlnGln   566

Qy   1325   AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGGAAG   1360
         ||| |||||||||||||||||||||
Db   567   LysAlaLeuGlnSerGlyProProGlnSerArgArgLys   579

RESULT 3
ABB75053
ID   ABB75053 standard; protein; 579 AA.
XX
AC   ABB75053;
XX
DT   01-MAY-2002 (first entry)
XX
DE   Human lung tumour L523S recombinant protein sequence SEQ ID NO:446.
XX
KW   Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW   immune response.
XX
OS   Homo sapiens.
XX
PN   WO200200174-A2.
XX
PD   03-JAN-2002.
XX
PF   28-JUN-2001; 2001WO-US021065.
XX
PR   28-JUN-2000; 2000US-00606421.

PR   02-AUG-2000; 2000US-00630940.
PR   21-AUG-2000; 2000US-00643597.
PR   15-SEP-2000; 2000US-00662786.
PR   09-OCT-2000; 2000US-00685696.
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI   Mcneill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
PI   Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX
DR   WPI; 2002-090513/12.
DR   N-PSDB; ABL49297.
XX
PT   Polynucleotides encoding lung tumor polypeptides, useful for treating
PT   lung cancer or stimulating an immune response.
XX
PS   Claim 2; Page 365-367; 374pp; English.
XX
CC   The present invention describes human lung tumour proteins. Human lung
CC   tumour proteins and polynucleotides have cytostatic and immunostimulant
CC   activities, and can be used in vaccine production. Compositions
CC   comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC   proteins, T cell populations, or antigen presenting cells that express
CC   the lung tumour proteins are useful for treating lung cancer or
CC   stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC   ABB75070 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 579 AA;

Alignment Scores:
Pred. No.:   2.36e-144   Length: 579
Score: 1639.00   Matches: 336
Percent Similarity: 83.89%   Conservative: 44
Best Local Similarity: 74.17%   Mismatches: 57
Query Match: 52.70%   Indels: 16
DB: 5   Gaps: 9

US-09-270-437D-5 (1-1708) x ABB75053 (1-579)

Qy   32 CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC 91
             |||||| ||::: |||| ||||||||
Db 133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152

Qy   92 TACATCCCGATGAGCAGATAGCA--CAGGGACCTGAGAATGGGCGCCAGGG------ 142
             |||||
Db 153 TyrIleProAspGluThrAlaAlaAlaGlnGlnAsnProLeuGluGlnProArgGlyArgArg 172

Qy 143 GGCTTTGGCTTCTCGGGGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCC 202
             |||||||::: ||:::
Db 173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-------GlySerValSer 189

Qy 203 AAGCAGCAGCAAGTGGACATCCCCCTTCGGCCTGCCACCCAGTATGTGGGTGCC 262
             |||||:::
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuProThrGlnPheGlnValGlyAla 209

Qy 263 ATTATTGGCAAGGAGGAGGGGCCACCATCCGCAACATCACAAAACAGCACCAGTCCAAGATA 322
             |||:::
Db 210 IleIleGluGlyLysGlyGlyAlaThrIleArgThrIleArgAsnIleThrLysGlnSerLysSerIle 229

Qy 323 GACGTGCCATAGGAAGGAGAACGCCAGGTGCAGCTGAAAAAGCCATCAGTGCACTCCACC 382
             |||||||::::::: ||:::::
Db 230 AspValHisArgLysGlyGluAsnAlaGlyAlaGlyIleAlaGluSerIleThrIleLeuSerThr 249

Qy 383 CCTGAGAGGCTGCTCCTCCGCTTGTAAGATATCTTGGAGATTATGCATAAAGAGGCTAAG 442
             |||||||:::
Db 250 ProGluGlyThrSerSerAlaAlaAlaCysLysSerSerIleLeuGluIleIleMetHisLysLysGluAlaGln 269

Qy 443 GACACCAAAAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCCCATAATAACTTTGTAGGG 502
         ||| |||||:::|||:::||||||||||||||

Db 270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly 289

Qy 503 CGTCTCATTGGCAAGGAGACGGAGAACCTGAAGAAGGTAGAGCAAGATACCGACAAAA 562
             |||||||||||
Db 290 ArgLeuIleGlyLysLysGluGlyLeuArgAsnLeuLysLysIleGluGlnAspThrAspThrLys 309

Qy 563 ATCACCATCCTCCTGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCACTGTG 622
             |||::: ||:::
Db 310 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

Qy 623 AAGGGGGGCCATCAGAATTGTTGCAGGACCAGGAAAATAATGACGAAAGTTCGGGAG 682
             ||||| ||::: ||:::|||
Db 330 LysGlyGlyAsnValGluThrCysAlaAlaLysLysIleMetLysLysIleArgGlu 349

Qy 683 GCCTATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCCTGAAC 736
             :::|||||| ||:::
Db 350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

Qy 737 CTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCG------CCT 787
             ||| :::||
Db 370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389

Qy 788 CCCAGCAGGGTTACTGGGGCTGCTGCCTATAGTCTCCTTATGCAGGCTCCCAGCAGGAG 847
             ||||| ::::
Db 390 ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu 406

Qy 848 ATGGTGCAGGTGTTTATCCCGCCAGGCAGTGGCGGCCATCATCGGCCAAGAAGGGGCAG 907
             ||::: |||||||:::||
Db 407 ThrValHisLeuPheIleProAlaLeuSerVal...GlyAlaIleLeuGlyLysGlnGlyGln 426

Qy 908 CACATCAAACAGCTTCCCCGGTTGCAGGCAGCGGCCTCCATCAGAGATTGCACCACCCGAAACA 967
         ||||||||| ::::|||
Db 427 HisIleLysSerArgPheAlaGlyAlaSerIleLysProLysIleAlaProAlaGluAlaAla 446

Qy 968 CCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCT 1027
             ||||| |||||
Db 447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466

Qy 1028 CAGGGAGAGATCTATGGCAAACTCAAGGAGGACGACTTCTTTGGTCCCAAGGAGGAAGTG 1087
             ||||||::: |||
Db 467 GlnGlyArgIleTyrGlyLysLeuLysGluAspAspPheValSerProLysGluGluVal 486

Qy 1088 AAGTTGGGACGACCCACATACGTGTGCCAGCATCAGCAGTTGCCGGGTCATTGCCAAGGT 1147
             |||:::   ::::
Db 487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506

Qy 1148 GGAAAAACCGTGAACGAGTTGCAGAATTGACCGAGTTGACGGAGTGGTAGTACCAAGAGAC 1207
             |||||:::
Db 507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerValAlaGluValValProArgAsp 526

Qy 1208 CAGACCCCTGATGAGAACCAGTCATCGTGAAAATCATCGGACATTTCTATGCCCAGT 1267
             ||||||::: |||:::
Db 527 GlnThrProAspGluAsnAspGlnValValLysIleIleThrGlyHisPheTyrAlaCys 546

Qy 1268 CAGATGGCTCAACGGAAGATCCCAGATCCAGTATGTGGGTGCC---CAGCATCAG 1324
             ||||||:::   ::::|||
Db 547 GlnValAlaGlnArgLysIleProAspIleGluLeuThrGlnValLysSerGlnHisGlnGln 566

Qy 1325 AAGGGA--CAGAGTAACCAGCGCCCAGGCACGGAGGAAG 1360
             ||||||:::
Db 567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 4
ABB74997
ID   ABB74997 standard; protein; 579 AA.
XX
AC   ABB74997;
XX
DT   01-MAY-2002 (first entry)
XX
DE   Human lung tumour L523S protein sequence SEQ ID NO:348.
XX
KW   Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;

```
KW  immune response.
XX
OS  Homo sapiens.
XX
PN  WO020200174-A2.
XX
PD  03-JAN-2002.
XX
PF  28-JUN-2001; 2001WO-US021065.
XX
PR  28-JUN-2000; 2000US-0606421.
PR  02-AUG-2000; 2000US-0630940.
PR  21-AUG-2000; 2000US-0643597.
PR  15-SEP-2000; 2000US-0662786.
PR  09-OCT-2000; 2000US-0685596.
PR  12-DEC-2000; 2000US-0735705.
PR  07-MAY-2001; 2001US-00850716.
XX
PA  (CORI-) CORIXA CORP.
XX
PI  Wang T,   Skeiky YAW,  Li SX,  Kalos MD,  Henderson RA;
PI  Mcneill PD,  Fanger N,  Retter MW,  Marnerakis M,  Fanger GR;
PI  Vedvick TS,  Carter D,  Watanabe Y,  Peckham DW;
XX
DR  WPI; 2002-090513/12.
DR  N-PSDB; ABL49254.
XX
PT  Polynucleotides encoding lung tumor polypeptides, useful for treating
PT  lung cancer or stimulating an immune response.
XX
PS  Example 2; Page 330-332; 374pp; English.
XX
CC  The present invention describes human lung tumour proteins. Human lung
CC  tumour proteins and polynucleotides have cytostatic and immunostimulant
CC  activities, and can be used in vaccine production. Compositions
CC  comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC  proteins, T cell populations, or antigen presenting cells that express
CC  the lung tumour proteins are useful for treating lung cancer or
CC  stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC  ABB75070 represent sequences used in the exemplification of the present
CC  invention
XX
SQ  Sequence 579 AA;

Alignment Scores:
Pred. No.:            2.36e-144
Score:               1639.00
Percent Similarity:    83.89%
Best Local Similarity: 74.17%
Query Match:           52.70%
DB:                  5

                                                   Length:        579
                                                   Matches:       336
                                                   Conservative:   44
                                                   Mismatches:     57
                                                   Indels:         16
                                                   Gaps:            9

US-09-270-437D-5 (1-1708) x ABB74997 (1-579)
```

```
Qy   32  CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCTGAAGGTCTCC   91
             ||||||   ||:::    |||||||
Db  133  ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluGlnPheThrLeuLysValAla  152

Qy   92  TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCCGAGGG------  142
Db  153  TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnAlaGlnGlnProArgGlyArgArg  172

Qy  143  GGCTTGGCTCCGGGGTCAGCCCCGCCAGGGGTCACCTGGTGCAGCAGGCCCCCCAGCC  202
             ||||||||||||
Db  173  GlyLeuGlyArgGlySerSerArgGlnGlySerPro-------GlySerValSer  189

Qy  203  AAGCAGCAGCAAGTGGACATCCCCTTCGGCTCCGGTGCCACCCAGTATGTGGGTGCC  262
             ||:::  ||||::::|||||:::||
Db  190  LysGlnLysProAspLeuLeuArgLeuuArgProLeuLeuValProThrGlnTyrGlnGlyGlyAla  209

Qy  263  ATTATTGCAAGAGGGGGCCACCATCCCAACATCACAAAACAGACCCCAGTCCAAGATA  322
             ||||||:::|||||||||||||
Db  210  IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrLysGlnThrGlnSerLysIle  229
```

```
Qy  323  GACGTGCATAGGAAGGAGGAGAACCCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC  382
             |||||||   |   |||||||   :::::    ||||||
Db  230  AspValHisArgLysGluAsnAlaGlyAlaAlaGlyAlaAlaGluLysProIleSerValThr  249

Qy  383  CCTGAGGGCTGCTCCTCCGCTGTAAGTGATCTTGAGATTATGCATAAAGAGGCTAAG  442
             |||||||:::  |||||||||:::||
Db  250  ProGluGlyThrSerAlaAlaCysLysSerAspLeuGluIleMetHisLysGluAlaGln  269

Qy  443  GACACCAAAAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCCCATAATAACTTTGTAGGG  502
             |||||||   :::   ||||||||
Db  270  AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly  289

Qy  503  CGTCTCATTGGCCAAGGAGGAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAA  562
             ||||||||||||:::   :::||||
Db  290  ArgLeuIleGlyLysGluGlyLysArgLeuLysLysLysIleGluGlnIleAspThrArgLys  309

Qy  563  ATCACCATCTCCTCGTTGCAAGACCTTACCCTTACAACCTGAGAGGACCATCACTGTG  622
             |||:::||||:::||||||||
Db  310  IleThrIleSerProLeuGlnLeuThrLeuTyrAsnProGluArgThrIleThrVal  329

Qy  623  AAGGGGCCATCGAGAATTGTTGCAGGGCCCGAGCAGCAGGAAATAATGAAGAAGTTCGGGAG  682
             ||||||:::   |||||:::||
Db  330  LysGlyAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu  349

Qy  683  GCCTATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCCTGAAC  736
             :::||||||:::|:::||||||
Db  350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn  369

Qy  737  CTGGCTGCTGTAGGTCTTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCG--------CCT  787
             ||||||   :::||
Db  370  LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro  389

Qy  788  CCCAGCAGCGTTACTGGGCGTGCTCCCTATAGCTCCTTATGCAGGCTCCCGAGCACGAG  847
             ||||||:::   |||   :::||
Db  390  ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu  406

Qy  848  ATGGTGCAGGTGTTATCCCCGCCAGCAGTGGGCGCCATCATCGGCAAGAAGAGGGGCAG  907
             |||   :::|:::   :::||
Db  407  ThrValHisLeuPheLeuProAlaLeuSerValGlyAlaIleIleGlyLysLysGlnGlyGln  426

Qy  908  CACATCAAACAGCTCTCCGGTTTGCCAGCGACGTTGGCAGGAAGCCTCATCAAGATTGCACCACCCGAAACCA  967
             |||||||||||   |||||||||||
Db  427  HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaAla  446

Qy  968  CCTGAGTTCCCAAAGTTCGTATGTTATCATCACTGGACCGCCAGAGGCCAATTCAAGGCT  1027
             ||||||:::||:::||
Db  447  ProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysAla  466

Qy  1028 CAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG  1087
             ||||||   :::|||||||
Db  467  GlnGlyArgIleTyrGlyLysLeuLysGluLysGluAsnPheValSerProLysGluGluVal  486

Qy  1088 AAGCTGGAGACCCCACATACGTGTGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGT  1147
             |||||||   :::|||
Db  487  LysLeuGluLysAspHisIleAspArgValProSerPheAlaAlaGlyAlaValIleGlyLysGly  506

Qy  1148 GGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGTGAGGTGGTAGTACCAAGAGAC  1207
             |||||||   :::||||||
Db  507  GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgArgAsp  526

Qy  1208 CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCCGGACATTTCTATGCCAGT  1267
             ||||||:::||||||   :::||
Db  527  GlnThrProAspGluAsnAspGlnValIleValLysIleIleArgThrPheLeuTyrArgAlaCys  546

Qy  1268 CAGATGGCTCAACGGAAGAATCCGAGACATCCTGGCCCAGTTAAGCAG---CAGCATCAG  1324
             ||||||:::   :::||||||
Db  547  GlnValAlaGlnArgLysAsnProArgHisProGlyProValLeuGlnHisGlnGlnGln  566

Qy  1325 AAGGGA---CAGAGTAACCAGCCCAGGCACGGAGGAAG  1360
             |||       |||||||||
Db  567  LysAlaLeuGlnSerGlyProProGlnSerArgArgLys  579
```

RESULT 5
ABB75054
ID   ABB75054 standard; protein; 579 AA.
XX
AC   ABB75054;
XX
DT   01-MAY-2002 (first entry)
XX
DE   Human lung tumour L523S recombinant protein sequence SEQ ID NO:449.
XX
KW   Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW   immune response.
XX
OS   Homo sapiens.
XX
PN   WO200200174-A2.
XX
PD   03-JAN-2002.
XX
PF   28-JUN-2001; 2001WO-US021065.
XX
PR   28-JUN-2000; 2000US-00606421.
PR   02-AUG-2000; 2000US-00630940.
PR   21-AUG-2000; 2000US-00643597.
PR   15-SEP-2000; 2000US-00662786.
PR   09-OCT-2000; 2000US-00685696.
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,   Skeiky YAW, Li SX, Kalos MD,   Henderson RA;
PI   Mcneill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
PI   Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX
DR   WPI; 2002-090513/12.
DR   N-PSDB; ABL49299.
XX
PT   Polynucleotides encoding lung tumor polypeptides, useful for treating
PT   lung cancer or stimulating an immune response.
XX
PS   Claim 2; Page 368-369; 374pp; English.
XX
CC   The present invention describes human lung tumour proteins. Human lung
CC   tumour proteins and polynucleotides have cytostatic and immunostimulant
CC   activities, and can be used in vaccine production. Compositions
CC   comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC   proteins, T cell populations, or antigen presenting cells that express
CC   the lung tumour proteins are useful for treating lung cancer or
CC   stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC   ABB75070 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 579 AA;

Alignment Scores:
Pred. No.:   2.36e-144
Score: 1639.00
Percent Similarity: 83.89%
Best Local Similarity: 74.17%
Query Match: 52.70%
DB: 5

Length: 579
Matches: 336
Conservative: 44
Mismatches: 57
Indels: 16
Gaps: 9

US-09-270-437D-5 (1-1708) x ABB75054 (1-579)

Qy   32 CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGTCTCC 91
          ||||||||||| ||| |||||||||| |||||||||||||||
Db   133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnThrPheLeuLysValAla 152

Qy   92 TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGGCCCGAGGG------ 142
          |||||| ||| ||| ||| ||| ||||
Db   153 TyrIleProAspGluThrValAsnGlnLeuAsnProLeuGlnProArgGlyArgArg 172

Qy   143 GGCTTGGCTCTCGGGTCAGCCCGGGGCTTCACCTGTGGCAGCGGGGGCCCCAGCC 202
          ||||| | ||| ||| ||| |||| :: :::
Db   173 GlyLeuGlyLeuArgGlySerSerArgGlnGlySerPro------GlySerValSer 189

Qy   203 AAGCAGCAGCAAGTGGACATCCCCTTGGCTCCTGGTGCCCACCAGTATGTGGGTGCC 262
          ||||:::|||||||||||||||||||
Db   190 LysGlnLysProCysAspLeuProLeuAlaProGlyAlaHisGlnTyrValGlyAla 209

Qy   263 ATTATTGGCAAGGAGGAGGGCCAACATCCGCAACTCACAAAACAGACCCAGTCCAAGATA 322
          |||||||||||| ||||||||||||||||||:|||||||||||
Db   210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229

Qy   323 GACGTGCATAGGAAGGAGGAACCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC 382
          |||||||||||||||||||||| ||||:||| |||||:::: |||||| |||
Db   230 AspValHisArgLysGluGluProGlyAlaAlaGluLysAlaIleSerValHisSerThr 249

Qy   383 CCTGAGGGGCTGCTCCTCCGGTTGAAGATGATCTTGGAAGATAATGCATAAAGAGGCTAAG 442
          ||||||||| |||||||::: |||||:::||||||||:::::::: |||||||
Db   250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluLIleMetHisLysGluAlaGln 269

Qy   443 GACACCAAAACGGTGACGAGGTTCCCTGAGACTCCTGGCCCATAATAACTTTGTAGGG 502
          |||||||||||||||||||||:::|||||||:::||||
Db   270 AspIleLysPheThrGluGluLProLeuLysIleLeuLeuAlaHisAsnAsnPheValGly 289

Qy   503 CGTTCATGGCAAGGAGGAGGACGAACCTGAGAAGGTAGAGCAAGATACCGACACAAAA 562
          |||||||||||||||||| ||||||||||||||||||||||||||:::|||
Db   290 ArgLeuIleGlyLysLysGluGlyLYArgAsnLeuLysSerGluGlnAspThrAspThrLys 309

Qy   563 ATCACCATCTCCTGTTGCAAGACCTACCCTTTACAACCTGAGAGACCCATCACTGTG 622
          ||| |||||||||||||||||||||:::||||| |||||||||||||||||:::
Db   310 IleThrIleSerProLeuGlnLeuThrLeuTyrAsnProGluGluThrIleThrVal 329

Qy   623 AAGGGGGCCCATCGAGAATTGTGCAGGGCGAGCAGCAGGAAATAATGAAGAAAGTTCGGGAG 682
          ||||| |||| ||::: ||||||||||||::||||||||||
Db   330 LysGlyAsnValGluThrCysAlaLysAlaGluGluIleIleLysMetLysLysIleArgGlu 349

Qy   683 GCCATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCTGGCCTGAAC 736
          ||||| | |||||||||||||::: |||||||||
Db   350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

Qy   737 CTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCG--------CCT 787
          ||||||||||::: |||||||||||:::
Db   370 LeuAsnAlaLeuGlyLeuPheProSerGlyMetProProProThrGlyPro 389

Qy   788 CCCAGCAGCGTTACTGGGGCTGCTCCCATAGCTCCTTTATCGAGGCTCCCAGCAGGAG 847
          ||||:::|||||||||||||::: ||||||||||:::
Db   390 ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer--GluThrGlu 406

Qy   848 ATGGTGCAGGTGTTTATCCCGGCCAGGCAGTGGCGCCATCATCGGCAAGAAGAGGGGCAG 907
          |||||||||| ||||| ||||||||||||||||||||:::||||||
Db   407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGln 426

Qy   908 CACATCAAACAGCTCTCCCGGTTGCCAGCGCCTCCATCAAGATTGCACACCCGAAACA 967
          |||||||||||||||||||||::: ||||||||||||
Db   427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysLysAlaProAlaGluAla 446

Qy   968 CCTGACTCCAAGTTCGTATGGTTATCATCACTGGACCGCAGAGGCCCAATTCAAGGCT 1027
          ||||||||||||:::||||||||::: |||||||||||
Db   447 ProAspAlaLysValAlaArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 466

Qy   1028 CAGGGAAGAATCATGGCAAACTCCAAGGAGGAGGAGACTTCTTCGTCCCAAGGAGGAAGTG 1087
          ||||||||||||::: |||||| |||||||||||
Db   467 GlnGlyArgIleTyrGlyTyrLeuLysIleLysGluGlnAspThrGlnGluGluLysVal 486

Qy   1088 AAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGTGGCCGGGTCATTGGCAAAAGT 1147
          |||||||||||| ||| |||||||||||||
Db   487 LysLeuGluAlaHisIleArgValProSerPheAlaLeuGlyArgValIleGlyLysGly 506

Qy   1148 GGAAAAACCGTGAACGAGTTGCAGAATTTGACGGCAGTGGAGTGGTAGTACCAAGAGAC 1207
          ||||||||||||||||::: ||||||| |||||
Db   507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgGlyArgAsp 526

Qy   1208 CAGACCCCTGATGAGAACGACCAGTCATCGTGAAAATCATCGGACATTTCTATGCCAGT 1267

```
Best Local Similarity:   74.17%        Mismatches:   57
Query Match:             52.70%        Indels:       16
DB:                      5             Gaps:          9

US-09-270-437D-5 (1-1708) x ABP61917 (1-579)

Qy    32  CGGGGAGCCATCATGAAGCTGAATGGCACCAGTTGAGAGAACCATGCCTGAAGGTCTCC   91
          |||||||||||:::|||||||||||||||||||||||||||||||||
Db   133  ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGlnAsnPheThrLeuLysValAla  152

Qy    92  TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAGTGGGCGCCAGGG------   142
          ||||:::||||||||||||||||   ||||||||||||||||||||| :::
Db   153  TyrIleProAspGluThrAlaAlaAlaGlnArgProLeuGlnLysProArgGlyArgArg  172

Qy   143  GGCTTTGGCTCTCGGGGTCAGCCCCCCAGGGTCACCTGTGGCAGCGGGGCCCCAGCC    202
          ||||:::|||||||||||||||||||||||||||||||||||   :::
Db   173  GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-------GlySerValSer  189

Qy   203  AAGCAGCAGCAAGTGGACATCCCCTTCGGCTCTCGTGCCCCAGTATGTGGGTGCC      262
          ||||||::::||||||||||||::|||||||:||||:||||||:::||||||||
Db   190  LysGlnLysProCysAspLeuProLeuArgLeuLeuLysValProThrGlnPheValGlyAla  209

Qy   263  ATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATA  322
          |||||||||||||||||||||||||||:||||||||||||||||||||||||
Db   210  IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle  229

Qy   323  GACGTGCATAGGAAGGAGAAGCGAGGTGCAGCTGAACATCAGTGTGCACTCCACC      382
          |||||||||||:||||||::::||||||||||||||:::|||||||:::
Db   230  AspValHisArgLysGluLysArgGlyAlaAlaAlaGlyGluGluIleThrIleLeuSerThr  249

Qy   383  CCTGAGGGCTGCTCCTCCGGTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG  442
          |||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db   250  ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln  269

Qy   443  GACACCAAAACCGGCTGACGAGGTTCCCTGAGAATCCTGGCCCAATAACTTTGTAGGG    502
          |||||||:|:::||||||||||||||||:||||::||||||||||::||||
Db   270  AspIleLysPheThrGluGlnIleProLeuLysIleLeuAlaHisAsnAsnPheValGly  289

Qy   503  CGTCTCATTGGCAAGGAGGACGGAACCTGAAGAAGTAGACAAGATACCGAGACAAAA     562
          ||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db   290  ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleLysIleProGluGlnLys  309

Qy   563  ATCACCATCTCCTCGTTGCAAGACTTACCCTTACAACCCTGAGAGGACCATCACTGTG    622
          ||||||||||||::||||||||||||:::|||||||||||||||||||||||||
Db   310  IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal  329

Qy   623  AAGGGGGCCATCGAGAATTGTGTGCGAGGGCCGAGCAGGAAATAATGAAGAAGTTCGGGAG  682
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   330  LysGlyAlaIleGluThrCysAlaLysAlaGlyAlaGluGlyIleAsnLysArgGlu    349

Qy   683  GCCTATGAGAATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGCCTGAAC  736
          :::||||||||||||:||||:||||||||||     :::||||||||||||||||||
Db   350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn  369

Qy   737  CTGGCTGCTGTAGGTCTTTTCCCAGTTCATCCAGCGCAGTCCCGCG------CCT     787
          ||||||||||||||:::|||||||||||||||||||||||||||   ||
Db   370  LeuAlaAlaLeuGlyLeuPheProThrSerGlyMetProProProThrGlyPro  389

Qy   788  CCCAGCGGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAG   847
          ||||:::||||:::||||||||||||||||||||||||||||||||||||
Db   390  ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu  406

Qy   848  ATGGTGCAGGTGTTTATCCCGCCGCCATCATCGGCAAGAAGGGGCAG      907
          ||||:::|||||||||||||||||   :::||||||||||||||||||||||
Db   407  ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleGlyLysLysGlnGln  426

Qy   908  CACATCAAACAGCTCTCCCGGTTGCCAGCGCCTCCATCAGCATTCCAGATTGCACCACCGCAAACA  967
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   427  HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla  446

Qy   968  CCTGACTCCAAAGTTCGTATGGTTATCATCATCGGACCGGCCAGAGCGCCATTCAAGGCT  1027
          ||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db   447  ProAspAlaLysValArgMetValIleIleIleGlyProAlaArgAlaProPheLysAla  466

Db   527  GlnThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCys  546
          |||||||||||:::|||||||  |||||||||||:|:|||||||||||||||||||
Qy  1268  CAGATGGCTCAACGGAAGATCCTGGACACATCCTGGCCCAGTTAAGCAG---CAGCATCAG 1324
          |||:::|||||:::||||||||||||||||:::|||||||   |||  |||
Db   547  GlnValAlaGlnArgLysIleLeuGluThrGlnIleLeuGlnHisGlnHisGlnGln  566

Qy  1325  AACGGA---CAGAGTAACCAGGGCCACGGCACGGAGGAAG 1360
          ||||||   ||||||:::||||||||||::|||||||||||
Db   567  LysAlaLeuGlnSerGlyProGlnSerArgArgLys  579
```

RESULT 6

```
ABP61917
ID   ABP61917 standard; protein; 579 AA.
XX
AC   ABP61917;
XX
DT   07-OCT-2002 (first entry)
XX
DE   Human lung cancer associated protein sequence SEQ ID NO:348.
XX
KW   Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200247534-A2.
XX
PD   20-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-US047576.
XX
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
PR   28-JUN-2001; 2001US-00897778.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI   Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI   Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR   WPI; 2002-583465/62.
DR   N-PSDB; ABQ92440.
XX
PT   Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT   the polynucleotides, useful in pharmaceutical compositions such as
PT   vaccines and as markers to indicate the presence of lung cancer.
XX
PS   Example 2; Page 337-339; 381pp; English.
XX
CC   The present invention describes isolated human lung carcinoma
CC   polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC   activity, and can be used in gene therapy and in vaccines. Compositions
CC   comprising (I) or (II) can be used for stimulating an immune response in
CC   a patient and for treating lung cancer in a patient. Oligonucleotides of
CC   (I) can be used for detecting the presence of a cancer in a patient, by
CC   obtaining a biological sample from the patient, contacting the biological
CC   sample with the oligonucleotide, detecting in the sample, an amount of
CC   polynucleotide that hybridises to the oligonucleotide and comparing the
CC   amount of polynucleotide that hybridises to the oligonucleotide to a
CC   predetermined cut-off value, and determining the presence of a cancer in
CC   the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC   vaccines. (I) is useful as a marker to indicate the presence or absence
CC   of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC   ABP61992 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 579 AA;

Alignment Scores:
Pred. No.:            2.36e-144
Score:               1639.00
Percent Similarity:  83.89%
Length:              579
Matches:             336
Conservative:        44
```

```
Qy  1028  CAGGGGAGGAATTATGCGAACTCAGGAGGAGGAACTTCTTGGTCCAGGAGGAAGTG 1087
          |||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db   467  GlnGlyArgIleTyrGlyLysIleLysGluGluLeuLeuValProLysGluGluVal 486

Qy  1088  AAGCTGGAGACCCACATACGTGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAGGT 1147
          ||||||||||| |||||||||::::  |||||||||||||||||||||||||
Db   487  LysLeuGluAlaHisIleArgProSerPheAlaAlaGlyArgValIleGlyLysGly 506

Qy  1148  GGAAAAACGGTGAACGAGTTCGAGAATTTGACGCGACGAGTCGTAGTACCAAGAGAC 1207
          ||||||||||||||||||||:::::||||||||:::|||||||||||:||||||
Db   507  GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526

Qy  1208  CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGT 1267
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   527  GlnThrProAspGluAsnAspGlnValIleValLysIleIleThrGlyHisPheTyrAlaCys 546

Qy  1268  CAGATGGCTCACGGAAGATCCGAGACATCCTGGCCCCAGGTTAAGCAG--CAGCATCAG 1324
          ||:::|||||||||||||:::|||||||||||:::|||||||||||||
Db   547  GlnValAlaGlnHisArgLysIleGluGlnIleLeuLeuThrGlnValLysGlnHisHisGlnGln 566

Qy  1325  AAGGGA--CAGAGTAACCAGGTCCCCAGGCACGGAGGAAG 1360
          |||     ||:::|||||||||||||||||||||||
Db   567  LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579
```

RESULT 7
ABP61974

ID   ABP61974 standard; protein; 579 AA.
XX
AC   ABP61974;
XX
DT   07-OCT-2002 (first entry)
XX
DE   Human lung cancer associated protein sequence SEQ ID NO:449.
XX
KW   Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200247534-A2.
XX
PD   20-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-US047576.
XX
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
PR   28-JUN-2001; 2001US-00897778.
XX
DR   WPI; 2002-583465/62.
DR   N-PSDB; ABQ92485.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI   Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI   Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR   WPI; 2002-583465/62.
DR   N-PSDB; ABQ92485.
XX
PT   Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT   the polynucleotides, useful in pharmaceutical compositions such as
PT   vaccines and as markers to indicate the presence of lung cancer.
XX
PS   Claim 9; Page 375-377; 381pp; English.
XX
CC   The present invention describes isolated human lung carcinoma
CC   polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC   activity, and can be used in gene therapy and in vaccines. Compositions
CC   comprising (I) or (II) can be used for stimulating an immune response in
CC   a patient and for treating lung cancer in a patient. Oligonucleotides of
CC   (I) can be used for detecting the presence of a cancer in a patient, by
CC   obtaining a biological sample from the patient, contacting the biological
CC   sample with the oligonucleotide, detecting in the sample, an amount of
CC   polynucleotide that hybridises to the oligonucleotide and comparing the
CC   amount of polynucleotide that hybridises to the oligonucleotide to a
CC   predetermined cut-off value, and determining the presence of a cancer in
CC   the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC   vaccines. (I) is useful as a marker to indicate the presence or absence
CC   of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC   ABP61992 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 579 AA;

Alignment Scores:
Pred. No.:   2.36e-144   Length: 579
Score: 1639.00   Matches: 336
Percent Similarity: 83.89%   Conservative: 44
Best Local Similarity: 74.17%   Mismatches: 57
Query Match: 52.70%   Indels: 16
DB: 5   Gaps: 9

US-09-270-437D-5 (1-1708) x ABP61974 (1-579)

```
Qy    32  CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGAGAACCATGCCTGAAGGTCC  91
          |||:::                                        :::
Db   133  ArgGlnAlaLeuAspLysProPheGlnLeuGluAsnGlyPheThrLeuLysValAla 152

Qy    92  TACATCCCGATGAGCAGATAGCA---CAGGACCTGAGAATGGCGCCGAGG------- 142
          ||||||||
Db   153  TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnLeuGlnProArgGlyArgArg 172

Qy   143  GGCTTGGCTCCTCGGCGTCAGCCCGCCGCCCCCAGGGTCACCTGTGGCAGCGGGGGCCCCAGCC 202
          |||||||:::                              |||       :::
Db   173  GlyLeuGlyLeuArgGlySerSerArgGlySerPro--------GlySerValSer 189

Qy   203  AAGCAGCAGCAAGTGGACATCCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCC 262
          |||:::                         |||
Db   190  LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnTyrValGlyAla 209

Qy   263  ATTATTGGACAAGGAGGGGGGCCACCATCCCAACATCACAAAACAGACCCAGTCCAAGATA 322
          |||                      :::
Db   210  IleIleGlyLysLysGlyLysAlaThrIleArgAsnIleThrArgLysGlnThrGlnSerLysIle 229

Qy   323  GACTGCAATAGGAAGGAAGAACGCAGGTGCAGTCAGTGTGAACTCCACC 382
          |||                   :::
Db   230  AspValHisArgLysGluAsnAlaGlyAlaAlaGlyLysSerIleThrIleLeuSerThr 249

Qy   383  CCTGAGGGCTGCTCCTCCGGTTGAAGATGATCCTTGAAGATTATGCATAAAGAGGCTAAG 442
          |||:::                        |||
Db   250  ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269

Qy   443  GACACCAAAACGGCTGACGAGGTTCCCCTGACGAGTCTGCCCATAATAACTTGTAGGG 502
          |||                         :::
Db   270  AspIleLysPheThrGluGluValIleProLeuLeuProLeuThrArgLeuAlaHisAsnAsnPheValGly 289

Qy   503  CGTCTCATTGGCAAGGAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGGACAAAA 562
          |||:::                                  |||
Db   290  ArgLeuIleGlyLysGluGlyLysArgAsnLeuLysLysLysGluGlnAspThrArgThrLys 309

Qy   563  ATCACCATCCTCCTGTTGCAAGACCTTTACCCTTTACAACCCTGAGAGACCATCACTGTG 622
          |||                            :::
Db   310  IleThrIleSerProLeuGlnLeuThrLeuTyrArgAsnProGluArgThrIleThrVal 329

Qy   623  AAGGGGGCCATCAGAGAATTGTTGCAGGGCCGAGCAGGAAATAATGAAGAAAGTTCGGGAG 682
          |||:::                                    |||
Db   330  LysGlyAsnValGlnThrCysAlaLysAlaGluGluLeuMetLysLysLysIleArgGlu 349

Qy   683  GCCATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCTGCCTGGCTGAAC 736
          :::|||:::||::::||:::::                 ||||||||
Db   350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

Qy   737  CTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCCCG------CCT 787
          |||||                          :::||||
Db   370  LeuAsnAlaLeuGlyLeuPheProPheSerGlyMetProProThrSerGlyPro 389

Qy   788  CCCAGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGAGGCTCCCGAGCAGGAG 847
```

```
PT   the polynucleotides, useful in pharmaceutical compositions such as
PT   vaccines and as markers to indicate the presence of lung cancer.
XX
PS   Claim 9; Page 372-374; 381pp; English.
XX
CC   The present invention describes isolated human lung carcinoma
CC   polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC   activity, and can be used in gene therapy, and in vaccines. Compositions
CC   comprising (I) or (II) can be used for stimulating an immune response in
CC   a patient and for treating lung cancer in a patient. Oligonucleotides of
CC   (I) can be used for detecting the presence of a cancer in a patient, by
CC   obtaining a biological sample from the patient, contacting the biological
CC   sample with the oligonucleotide, detecting in the sample, an amount of
CC   polynucleotide that hybridises to the oligonucleotide and comparing the
CC   amount of polynucleotide that hybridises to the oligonucleotide to a
CC   predetermined cut-off value, and determining the presence of a cancer in
CC   the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC   vaccines. (I) is useful as a marker to indicate the presence or absence
CC   of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP1866 to
CC   ABP61992 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 579 AA;

Alignment Scores:
Pred. No.:
Score:                  2.36e-144      Length:        579
                        1639.00        Matches:       336
Percent Similarity:     83.89%         Conservative:  44
Best Local Similarity:  74.17%         Mismatches:    57
Query Match:            52.70%         Indels:        16
DB:                     5              Gaps:          9

US-09-270-437D-5 (1-1708) x ABP61973 (1-579)

Qy    32 CGGGGAGCCATCATGAAGCTGAATGCCACCAGTTGAGAGACCATGCCCTGAAGGTCTCC  91
            |||||||:::|| |||||||||    |||||||||||
Db   133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluThrLeuLysValValAla 152

Qy    92 TACATCCCGATGAGCCAGATAGCA---CAGGGACCTGAGAATGGGCGCCAGGG----- 142
            |||||||||||||||||||||||    ||||||||||||||||||||||||||
Db   153 TyrIleProAspGluThrAlaAlaAla-GlnGlnAsnProLeuGlnProArgGlyArgArg 172

Qy   143 GGCTTTGGCTCTCGGGTCAGCGTCCCCAGGGTCCACCTGTGGCAGGGGGGCCCCAGCC 202
            |||:|||| ||||||||||||||      |||      ||| :||  :::
Db   173 GlyLeuGlyGlnArgGlySerArgGlyGlySerPro--------GlySerValSer 189

Qy   203 AAGCAGCAGCAAGTGGACATGGACATCCCCCTTCGGCTCCGTGCCCCACCCAGTATGTGGGTGCC 262
            |||||||:  :||  :||||||||||||||||||||||||||:||:||||
Db   190 LysGlnLysProCysAspSerLeuProLeuLeuArgLeuArgValProThrGlnPheValGlyAla 209

Qy   263 ATTATTGGCAAGGAGGGGGGCCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATA 322
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   210 IleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleThrArgGlnThrGlnSerLysIle 229

Qy   323 GACGTGCATAGGAGGAAGGAGAACGCAGGTGCAGCTGCGAAAAAGGCCATCAGTGTGCACTCCCACC 382
            |||||||||||| |||||||||||:::|||:::||||||:::||||||:::||
Db   230 AspValHisArgLysGluAsnAlaGlyAlaAlaAlaGluLysSerSerIleThrLeuSerThr 249

Qy   383 CCTGAGGGGCTGCTCCTCCGGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG 442
            ||||||||||||||||||:::|| ||:::||||||||||||||||||||||||||||
Db   250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln 269

Qy   443 GACACCAAAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCATGAGATCCTGGCCCATAATAACTTTGTAGGG 502
            |||||||      ::  :|||||||||||||||:::||:::||:::|||
Db   270 AspIleLysPheThrGluGlyIleLeuProLeuLysIleLeuAlaHisAsnAsnPheValGly 289

Qy   503 CGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAA 562
            |||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db   290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysValGluGlnAspThrGluThrLys 309

Qy   563 ATCACCATCTCCTCGTTGCAAGACCTTACCCTTACAACCCTGAGACCCTGAGACCACCATCACTGTG 622
            ||||||||:::|| |||||:::||||:::||
Db   310 IleThrIleSerProLeuGlnAspLeuThrLeuGlnThrLeuArgThrIleThrVal 329
```

```
Db   390 ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer--GluThrGlu 406
            |||:::::|||          ||||||||||||          ||| |||
Qy   848 ATGGTGCAGGTGTTTATCCCCCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGGCAG 907

Db   407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysLysGlyGln 426
            :::||||||||||||||||||||||||||||||||||||||:::||||||||||||||
Qy   908 CACATCAAACAGCTCTCCGGTTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACA 967

Db   427 HisIleLysGlnLeuSerArgPheAlaGlyLeuAlaSerIleLysIleAlaProAlaGluAla 446
            |||||:::||||||||||||||||||||||||||||||||||||||||||
Qy   968 CCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGCCCAATTCAAGGCT 1027

Db   447 ProAspAlaLysValArgMetValIleIleThrGlyProProGluProAlaGlnPheLysAla 466
            |||||||:::||||||||||||||||||||||||||||||||||||||||||
Qy  1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCAAGGAGGAAGTG 1087

Db   467 GlnGlyArgIleTyrGlyLysLeuLysGluGluAsnPheValSerProLysGluGluVal 486
            |||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1088 AAGCTGGAGACCCACATACGTGCCAGCATCAGCAGCGGGGTCATTGGCAAAGGT 1147

Db   487 LysLeuGluAlaHisIleArgAlaSerIleSerSerGlyValIleGlyLysLysGly 506
            |||||||||||||||||||||||:::||
Qy  1148 GGAAAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGGTGTGTAGTACCAAGAGAC 1207

Db   507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgArgAsp 526
            |||||||||||||||||||||:::|||:::||||||||||||||||||||
Qy  1208 CAGACCCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGT 1267

Db   527 GlnThrProAspGluAspAspGlnValIleValLysIleIleGlyHisPheTyrAlaCys 546
            |||||||||||||||||||||:::||
Qy  1268 CAGATGGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAG---CAGCATCAG 1324

Db   547 GlnValAlaGlnArgLysIleGlnGlyIleLeuGluThrGlnValLysGlnHisGlnGln 566
            |||:::::|||:::||||||||:::||
Qy  1325 AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGGAAG 1360

Db   567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579
            |||:::|||||||||||||||:::||
Qy   579
```

```
RESULT 8
ABP61973
ID   ABP61973 standard; protein; 579 AA.
XX
AC   ABP61973;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Human lung cancer associated protein sequence SEQ ID NO:446.
XX
KW   Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
XX   WO200247534-A2.
PN
XX
PD   20-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-US047576.
XX
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
PR   28-JUN-2001; 2001US-00897778.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI   Mcneill PD, Fanger MW, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI   Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR   WPI; 2002-583465/62.
DR   N-PSDB; ABQ92483.
XX
PT   Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
```

```
Qy   623  AAGGGGGCCATCGAGAATTGTGCAGGCCGAGCAGGAAATAATGAAGAAAGTTCGGGAG  682
               |||  ::|  |||||||     ||| |:::|||  ::||||||||
Db   330  LysGlyAlaThrGluThrGluAlaGlyGluGluIleMetLysLysIleArgGlu       349

Qy   683  GCCTATGAGAATGATGGGCTGCCATGAGC------TCTCACCTGATCCTGGCTGAAC    736
              :::||||||||| ||| |||||                |||||  |||||||
Db   350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

Qy   737  CTGGCTGCTGTGTAGGTCTTTTCCCAGCTTCATCCAGCCGCAGTCCGCCG-------CCT 787
              |||  |||           :::||| |||      || |||||  ||
Db   370  LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro    389

Qy   788  CCCAGCAGCGTTACTGGGCGCTGCTCCCTATAGCTCCTTTATGCAGCCTCCCGAGCAGGAG 847
              ||||| |||              ||| ||||  |||||  |||||||
Db   390  ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu  406

Qy   848  ATGGTGCAGGTGTTTATCCCCGCCCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGCAG  907
              |||  |||||||||  ::|||||||        |||||||||||| |||:::|||||
Db   407  ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysLysGlnGln  426

Qy   908  CACATCAAACAGCTCTCCCGTTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACA   967
              ||||||||||||    ||  |||||||||||||||||||||  ||||
Db   427  HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla  446

Qy   968  CCTGACTCCAAGTTCGTATGTTATCATCACTGACCACCGGCCAGAGGCCCCAATTCAAGGCT 1027
              |||||||||||||  ||                     :::
Db   447  ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla  466

Qy  1028  CAGGGAGAATCATGCAAACTCAAGGACGGAGAACTTCTTTGGTCCAAGGAGGAAGTG    1087
              |||||||||          ||||  |   |||||||:::|||||||||||
Db   467  GlnGlyArgIleTyrGlyLysLysLysSerProLysSerGluGluVal             486

Qy  1088  AAGCTGAGACCCCACATACGTGCCAGCATCAGCAGCTCGCCGGGTCATTGGCAAAGGT   1147
              ||| ||||||||       ||||||||||||||  |||||||||:::|||||
Db   487  LysLeuGluLysHisIleArgValProSerPheAlaAlaGlyValIleGlyLysGly    506

Qy  1148  GGAAAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGAC 1207
              |||||  |||||||||||||||||       :::|||||||||||
Db   507  GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValValValProArgAsp 526

Qy  1208  CAGACCCCTGATGAGAACGAACGACCAGTCATCGTGAAAATCATCGACATTTCTATCGCCAGT 1267
              ||| ||||  |||||||||||:::||||||||||||||| |||
Db   527  GlnThrProAspGluAsnGluArgProValIleValLysIleIleThrGlyPheHisGlyTyrAlaCys 546

Qy  1268  CAGATGGCTCAACGGAAGATCCGAGAATCCTGGCCCAGGTTAAGCAG---CAGCATCAG  1324
              ||| |||||        ||||||        :::||||||||
Db   547  GlnValAlaGlnArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGln    566

Qy  1325  AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGGAAG                      1360
              |||          |||       |||||:::||||||
Db   567  LysAlaLeuGlnSerGlyProGlnSerGlnSerArgArgLys                   579
```

RESULT 9
ADA28536

```
ID   ADA28536   standard; protein; 579 AA.
XX
AC   ADA28536;
XX
DT   20-NOV-2003   (first entry)
XX
DE   Recombinant human lung tumour protein L523S #1.
XX
KW   cancer; lung cancer; gene therapy; vaccine; human;
KW   lung squamous cell carcinoma.
XX
OS   Homo sapiens.
XX
XX
PN   US2003064947-A1.
XX
PD   03-APR-2003.
XX
PF   30-NOV-2001; 2001US-00007700.
```

```
XX
PR   18-MAR-1998;   98US-00040802.
PR   27-JUL-1998;   98US-00123912.
PR   22-DEC-1998;   98US-00221107.
PR   02-APR-1999;   99US-00285479.
PR   17-DEC-1999;   99US-00476496.
PR   30-DEC-1999;   99US-00480884.
PR   10-JAN-2000;   2000US-00510376.
PR   22-FEB-2000;   2000US-00542615.
PR   04-APR-2000;   2000US-00606421.
PR   28-JUN-2000;   2000US-00630940.
PR   02-AUG-2000;   2000US-00643597.
PR   21-AUG-2000;   2000US-00662786.
PR   15-SEP-2000;   2000US-00685696.
PR   09-OCT-2000;   2000US-00735705.
PR   12-DEC-2000;   2001US-00850716.
PR   07-MAY-2001;   2001US-00897778.
PR   28-JUN-2001;
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI   Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI   Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR   WPI; 2003-540798/51.
XX
PT   New isolated polynucleotides and polypeptides useful for diagnosing,
     preventing and/or treating cancer, particularly lung cancer.
XX
PS   Claim 9; Page 285-287; 296pp; English.
XX
CC   The invention describes isolated polynucleotides and polypeptides useful
CC   for diagnosing, preventing and/or treating cancer, particularly lung
CC   cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC   defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
CC   specification; complements of the nucleotide sequences cited above; at
CC   least 10 contiguous residues of any of the nucleotide sequences cited above; a
CC   sequence that hybridise to any of the nucleotide sequences under highly
CC   stringent conditions; a sequence that is at least 75 or 90% identical to
CC   the above nucleotide sequences; or degenerate variants of the above
CC   nucleotide sequences. The composition and methods are useful in
CC   diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC   in gene therapy and in vaccines. This is the amino acid sequence of a
CC   recombinant human lung tumour associated protein.
XX
SQ   Sequence 579 AA;
```

Alignment Scores:

Pred. No.:	2.3e-144	Length:	579
Score:	1639.00	Matches:	336
Percent Similarity:	83.89%	Conservative:	44
Best Local Similarity:	74.17%	Mismatches:	57
Query Match:	52.70%	Indels:	16
DB:	7	Gaps:	9

US-09-270-437D-5 (1-1708) x ADA28536 (1-579)

```
Qy    32  CGGGGAGCCATCATGAAGCTGATGAGCTGGAGCACCAGTTGGAGAACCATGCCTGAAGGTCTCC  91
              |||      |||||:::   ||||||||||||||||  ||||||||||||||:::
Db   133  ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla    152

Qy    92  TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCCGCGAGGG------  142
              ||||||||||||||||||       |||||||||||||||||||||||
Db   153  TyrIleProAspGluThrAlaAlaGlnAlaGlnAsnProLeuGlnAsnProLeuArgGlyArg  172

Qy   143  GGCTTTGGCTCTCGGGGTCAGCCCCGCCAGGGCTCACCTGTGCAGCCCAGCGGGGCCCCAGCC  202
              ||  ||||||||  |||||||||||||||||     |||||||||:::||||||||||
Db   173  GlyLeuGluGlyGlnAlaGlyGlySerSerArgGlnGlySerPro-------GlySerValSer  189

Qy   203  AAGCAGCAGCAAGTGGACATCCCCTTCGGCATCCCCTGTGCCCACCCAGTATGTGGGTGCC    262
              |||:::      |||||||||||||||||||||||||||||:::|||||||
Db   190  LysGlnLysProCysAspLeuProLeuArgLeuArgLeuValProThrGlnPheGlnPheValGlyAla 209
```

```
Qy    1325 AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGGAAG 1360
           |||   |||||||:::|||||::|||||||||
Db     567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 10
ADA28539
ID   ADA28539 standard; protein; 579 AA.
XX
AC   ADA28539;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Recombinant human lung tumour protein L523S #2.
XX
KW   cancer; lung cancer; gene therapy; vaccine; human;
KW   lung squamous cell carcinoma.
XX
OS   Homo sapiens.
XX
PN   US2003064947-A1.
XX
PD   03-APR-2003.
XX
PF   30-NOV-2001; 2001US-00007700.
XX
PR   18-MAR-1998;    98US-00040802.
PR   27-JUL-1998;    98US-00123912.
PR   22-DEC-1998;    98US-00221107.
PR   02-APR-1999;    99US-00285479.
PR   17-DEC-1999;    99US-00465396.
PR   30-DEC-1999;    99US-00476496.
PR   10-JAN-2000;    2000US-00480884.
PR   22-FEB-2000;    2000US-00510376.
PR   04-APR-2000;    2000US-00542615.
PR   28-JUN-2000;    2000US-00606421.
PR   02-AUG-2000;    2000US-00630940.
PR   21-AUG-2000;    2000US-00643597.
PR   15-SEP-2000;    2000US-00662786.
PR   09-OCT-2000;    2000US-00685696.
PR   12-DEC-2000;    2000US-00735705.
PR   07-MAY-2001;    2001US-00850716.
PR   28-JUN-2001;    2001US-00897778.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,    Wang A,    Skeiky YAW,    Li SX,    Kalos MD,    Henderson RA;
PI   Mcneill PD,    Fanger N,    Retter MW,    Durham M,    Fanger GR,    Vedvick TS;
PI   Carter D,    Watanabe Y,    Peckham DW,    Cai F,    Foy TM;
XX
DR   WPI; 2003-540798/51.
XX
PT   New isolated polynucleotides and polypeptides useful for diagnosing,
PT   preventing and/or treating cancer, particularly lung cancer.
XX
PS   Claim 9; Page 288-289; 296pp; English.
XX
CC   The invention describes isolated polynucleotides and polypeptides useful
CC   for diagnosing, preventing and/or treating cancer, particularly lung
CC   cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC   defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
CC   specification; complements of the nucleotide sequences cited above; at
CC   least 10 contiguous residues of the nucleotide sequences cited above; a
CC   sequence that hybridise to any of the nucleotide sequences under highly
CC   stringent conditions; a sequence that is at least 75 or 90% identical to
CC   the above nucleotide sequences; or degenerate variants of the above
CC   nucleotide sequences. The composition and methods are useful in
CC   diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC   in gene therapy and in vaccines. This is the amino acid sequence of a
CC   recombinant human lung tumour associated protein.
XX
SQ   Sequence 579 AA;

Alignment Scores:
```

```
Qy    263 ATTATTGGCAAGGAGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATA 322
          ||||| || ||||| ||||||||||||||||||||||||||   |||||||||||||||
Db    210 IleIleGlyLysGluGlyGlyHisHisProGlnHisHisLysThrAspProSerLysIle 229

Qy    323 GACGTGCATAGGAGAGGGAGAACGCCAGGTGCAGCTGAAAAAGCCATCAGTGTGCCACTCCACC 382
          |||  |   |||      |||||:::|||::|||||||||
Db    230 AspValHisArgGlyGluAsnAlaGlyAlaAlaGluLysAlaIleSerValProLeuSerThr 249

Qy    383 CCTGAGGGGCTGCCTCCCGCTTGTGAAGATGATCTTGGAGATTATGCATAAAGAGGTAAG 442
          ||||| ||||   |||||||||||||||::||:|||||||||||
Db    250 ProGluGlyThrSerAlaAlaCysLysSerLeuGluLeuMetHisLysGluAlaGln 269

Qy    443 GACACCAAAACGGCTGACGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGG 502
          ||| ||||     :: |||||||||||||||||::||:|||||||||
Db    270 AspIleLysPheGluThrGlyLeuLysLysIleLeuAlaHisAsnAsnPheValGly 289

Qy    503 CGTCTCATTGGCAAGGAAGGACGGAACCTGAGAAGGTAGAGCAAGATACCGAGACAAAA 562
          ||| |||| ||||||||||||:::|||||||||||||||
Db    290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysGluGlnAspThrAspThrLys 309

Qy    563 ATCACCATCTCCGTTGCAAGACCTTACCCTTTACACCCTGAGAGGACCACTCACTGTG 622
          |||||||  ||::|||||||||||||||::|||||
Db    310 IleThrIleSerProLeuGlnLeuThrLeuTyrAsnProGluArgThrThrVal 329

Qy    623 AAGGGGGCCATCGAGAATTGTTGCAGGGGCCAGGCAGGAAAATAATGAAGAAAGTTCGGGAG 682
          ||| || ||||||||||||||||||::|||||||||||
Db    330 LysGlyAlaAsnValGluThrCysAlaLysAlaGluGluIleIleMetLysLysIleArgGlu 349

Qy    683 GCCTATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCTGAAC 736
          ||::||:   |||||||||||||||||||      |||||||||
Db    350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

Qy    737 CTGGCTGCTGTGTAGGTCTTTTCCCAGCTTCATCCAGCCAGTCCGCCG-------CCT 787
          ||||||| ||::|||||||||||||      ||||
Db    370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389

Qy    788 CCCAGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAG 847
          |||||||||||||||::|||||||||||::|||
Db    390 ProSerAlaMetThr------ProProTyrProGlnPheGlnSer---GluThrGlu 406

Qy    848 ATGGTGCAGGTGTTATCCCCGCCAGCAGTGGGCGGCCATCATCGGCAAGAAGAGGGCAG 907
          |||||||||| |   :::||||||||||||||||||||||||
Db    407 ThrValHisLeuPheIleProArgAlaLeuSerValGlyAlaIleIleGlyLysLysArg 426

Qy    908 CACATCAAACAGCTCTCCGGTTTGCCAGGCGCCTCCATCAAGATTGCACCACCCGAAACA 967
          ||||||||||||::| ||||||||||||||||||||||||
Db    427 HisIleLysGlnLeuSerArgPheAlaGlyArgLeuAlaSerIleLysIleAlaProAlaGluAla 446

Qy    968 CCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCGCCAGAGGCCCAATTCAAGGCT 1027
          ||||||||||||||||  |:::||||||||||||||||||
Db    447 ProAspAlaLysValAlaArgMetValIleIleThrGlyProProGlnIlePheLysAla 466

Qy   1028 CAGGGAGAAATCATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCAAGGAGGAAGTG 1087
          |||| ||::|||||||||||||||||||||||||||||
Db    467 GlnGlyArgIleTyrGlyIleTyrArgGlyIleLeuLysGluLysGluGlyValSerProLysGlyGluVal 486

Qy   1088 AAGCTGGAGACCCACATTACCGTGTGCCAGGCAGCTGGCCCGGGTCATTGGCAAAGGT 1147
          ||||||||||||||||   :|||||||||||||||||||
Db    487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysLysGly 506

Qy   1148 GGAAAAACGGTGACGAGTTGACAGAATTTGACCCAGCTGAGTGGTAGTACCCAAGAGAC 1207
          ||| ||||||||||||||||||::|||||||||||||||
Db    507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgArgAsp 526

Qy   1208 CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCATGCCAGT 1267
          |||||||||  | ||::|||||||||||||||||||||
Db    527 GlnThrProAspGluAsnAspGlnValIleValLysIleIleGlyHisPheMetProVal 546

Qy   1268 CAGATGGCTCAACGGAGATCCGAGACATCCTGGCCCCAGGTTAAGCAG---CAGCATCAG 1324
          |||::||||||  |||||||||||||||::|||||
Db    547 GlnMetAlaGlnArgArgLysIleLeuGluLeuLeuThrGlnIleValLysGlnHisGlnGln 566
```

Pred. No.:
Score: 2.36e-144
Percent Similarity: 1639.00
Best Local Similarity: 83.89%
Query Match: 74.17%
DB: 52.70%

Length: 579
Matches: 336
Conservative: 44
Mismatches: 57
Indels: 16
Gaps: 9

US-09-270-437D-5 (1-1708) x ADA28539 (1-579)

```
Qy    32 CGGGGAGCCATCATGAAGTGAATGGCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC 91
            ||||||     ||||||||     :::||||
Db   133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152

Qy    92 TACATCCCGATGAGCAGATAGCA---CAGGGACTTGAGAATGGGCGCCGAGGG------ 142
            ||||||        |||      ||||
Db   153 TyrIleProAspGluThrAlaAlaGlnProLeuGlnProArgGlnProArgArgArg 172

Qy   143 GGCTTTGGCTCTCGACGGGTCAGCCCCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCC 202
            |||||                          ||||||
Db   173 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro--------GlySerValSer 189

Qy   203 AAGCAGCAGCAAGTGGACATCCCCTTCGGCTCCTGTGCCCACCCAGTATGTGGGTGCC 262
            |||||   :::                         ||||||
Db   190 LysGlnLysProCysAspSerLeuProLeuLeuValProThrGlnPheValGlyAla 209

Qy   263 ATTATTGGCAAGGAGGGGCCACCATCCGACAACATCACAAAAACAGACCCAGTCCAAGATA 322
            ||||||||                     |||  |||  :::  ||||||
Db   210 IleIleGlyLysGluGlyGluGln--AlaThrIleArgAsnIleThrLysGlnIleHisThrLysGlnSerIle 229

Qy   323 GACGTGCATAGGACAGGAGAAGGAGAACCCAGTGCAGCTGGAAAAAAGCCATCAGTGTGCACTCCACC 382
            |||||||                     ||  :::||||||     ||||||
Db   230 AspValHisArgLysGluAlaGluAlaLysAlaAlaGluLeuLysSerIleThrLeuSerThr 249

Qy   383 CCTGAGGGGCTGCTCCTCCGCTTGTGAAGATCTTGGATTATGCATAAAGAGAGGCTAAG 442
            |||||                     :::||||||
Db   250 ProGluGlyThrSerAlaAlaAlaCysLysLysSerIleLeuGluLeuMetHisLysGlnGln 269

Qy   443 GACACCAAAACGGCTGACGAGTTCCCCTGAAGATCCTGGCCCCATAATAACTTGTAGGG 502
            ||||     :::|||||                     ||||||
Db   270 AspIleLysPheThrGluThrAlaAspGluPheProLeuLysIleLeuProLeuLeuAlaHisAsnAsnPheValGly 289

Qy   503 CGTCTCATTGGCAAGGAAGGACGGAACCTGAGGAGGTAGAGCAAGATACCGAGACAAAA 562
            ||||||     :::|||||
Db   290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysSerIleLysSerIleGluGlnAspThrAspThrLys 309

Qy   563 ATCACCATCTCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGGAGGACCATCACTGTG 622
            ||||||                     ||||||
Db   310 IleThrIleSerProLeuGlnLeuThrLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

Qy   623 AAGGGGGGCCATCGAGAATTGTTGCAGGGCCCAGCAGGAAAATAATGAAGAAGTTCGGGAG 682
            ||||||     :::|||||
Db   330 LysGlyValAsnValGluThrCysAlaAlaLysGluGluIleMetLysLysIleArgGlu 349

Qy   683 GCCTATCGAGAATGATGGCTGCCATGAGC-----TCTCACCTGATCCTGGCCTGAAC 736
            :::||||||     :::|||||
Db   350 SerTyrGluGluAspIleAlaSerMetAsnLeuGlnAlaHisLeuProGlyLeuAsn 369

Qy   737 CTGGCTGCTGTAGGTCTTTTCCAGTTCATCCAGCAGTCCCGCCG--------CCT 787
            ||||||                     ||||||
Db   370 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 389

Qy   788 CCCAGCAGCGGTTACTGGGGCTGCTCCCTATAGTCTCCTTTATGCAGGCTCCGAGCAGGAG 847
            |||||     :::||||||
Db   390 ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu 406

Qy   848 ATGGTGCAGGTGTTTATCCCCGCCAGGCAGTGGCGCCATCATCGGCAAGAAGGGGCAG 907
            ||||||     ||:::|||||
Db   407 ThrValHisLeuPheIleProAlaLeuSerValAlaIleIleGlyLysGlnGlyGln 426

Qy   908 CACATCAAACAGCTCTCCCGGTTTGCCAGCCGCCCTTCCATCAAGAGATTGCACCACCCGAAACA 967
            ||||||     :::|||||
Db   427 HisIleLysGlnLeuSerArgPheAlaSerArgProSerIleLysArgLeuHisHisProLysThr 446
```

```
Qy   968 CCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACGCGCAGAGGCCCAATTCAAGGCT 1027
            ||||||     :::|||||
Db   447 ProAspAlaLysValArgMetValIleIleThrGlyProProAlaGlnPheLysAla 466

Qy  1028 CAGGGAAGAATCATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG 1087
            ||||||     :::||||
Db   467 GlnGlyArgIleTyrGlyLysLysIleLeuGluAsnPheValSerProLysGluGluVal 486

Qy  1088 AAGTTGGAGACCCACATACGTGTGCCAGATACAGCAGTCTGGCGCGGGTCATTGGCAAAGGT 1147
            ||||||     ||||||
Db   487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyValIleGlyLysGly 506

Qy  1148 GGAAAAACGGTGAACGAGTTGCAGAATTTGACGGTGGTAGTACCAAGAGAC 1207
            ||||||     :::||||
Db   507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp 526

Qy  1208 CAGACCCCTGATGAGAAGCAACCAGGTCATCGTGAAAATCATCGGACATTTCATGCCAGT 1267
            ||||||     :::||||
Db   527 GlnThrProAspGluAsnGlnProGlyHisArgGluAsnHisArgThrGlyHisPheTyrAlaCys 546

Qy  1268 CAGATGGCTCAACCGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAG---CAGCATCAG 1324
            ||||||     :::|||||
Db   547 GlnValAlaGlnArgLysIleGlyIleLeuLeuThrGlnValLysGlnHisGlnGln 566

Qy  1325 AAGGGA--CAGAGTAACCAGGCCCAGGCACGGAGGAAG 1360
            |||||
Db   567 LysAlaLeuGlnSerGlyProGlnSerArgArgLys 579
```

RESULT 11
ADA28438
ID ADA28438 standard; protein; 579 AA.
XX
AC ADA28438;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human lung tumour protein L523S.
XX
KW cancer; lung cancer; gene therapy; vaccine; human;
KW lung squamous cell carcinoma.
XX
OS Homo sapiens.
XX
PN US2003064947-A1.
XX
PD 03-APR-2003.
XX
PF 30-NOV-2001; 2001US-00007700.
XX
PR 18-MAR-1998; 98US-00040802.
PR 27-JUL-1998; 98US-00123912.
PR 22-DEC-1998; 98US-00221107.
PR 02-APR-1999; 99US-00285479.
PR 17-DEC-1999; 99US-00463396.
PR 30-DEC-1999; 99US-00476495.
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510375.
PR 04-APR-2000; 2000US-00542615.
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685696.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX

DR   WPI; 2003-540798/51.
DR   N-PSDB; ADA28437.
XX
PT   New isolated polynucleotides and polypeptides useful for diagnosing,
PT   preventing and/or treating cancer, particularly lung cancer.
XX
PS   Example 2; Page 249-251; 296pp; English.
XX
CC   The invention describes isolated polynucleotides and polypeptides useful
CC   for diagnosing, preventing and/or treating cancer, particularly lung
CC   cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC   defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
CC   specification; complements of the nucleotide sequences cited above; at
CC   least 10 contiguous residues of the nucleotide sequences cited above; a
CC   sequence that hybridise to any of the nucleotide sequences under highly
CC   stringent conditions; a sequence that is at least 75 or 90% identical to
CC   the above nucleotide sequences; or degenerate variants of the above
CC   nucleotide sequences. The composition and methods are useful in
CC   diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC   in gene therapy and in vaccines. This is the amino acid sequence encoded
CC   by a human lung tumour cDNA isolated from a lung squamous cell carcinoma
CC   that may be useful in the diagnosis and treatment of lung cancer and
CC   other disorders.
XX
SQ   Sequence 579 AA;

Alignment Scores:
Pred. No.:   2.36e-144   Length: 579
Score: 1639.00   Matches: 336
Percent Similarity: 83.89%   Conservative: 44
Best Local Similarity: 74.17%   Mismatches: 57
Query Match: 52.70%   Indels: 16
DB: 7   Gaps: 9

US-09-270-437D-5 (1-1708) x ADA28438 (1-579)

Qy   32 CGGGGACCATCATGAAGTCGAATGGCCACAGTTGAGAACCATGCCTGAAGGTCTCC 91
                 ||||::: ||| ||||| ||||| |||||||||||| ||||||
Db   133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGlnLeuAsnPheThrLeuLysValAla 152

Qy   92 TACATCCCGATGACGATAGCA---CAGGGACCTGAGAATGGGCCGCCGAGGG------ 142
                 |||||||||| ||| |||||| ||
Db   153 TyrIleProAspGluThrAlaAlaAlaGlnThrAlaAlaGlnIleAsnProLeuGlnProLeuGlnProArgGlyArgArg 172

Qy   143 GGCTTTGGCTCTCGGGGTCAGCCCCCCAGGGGCGTCACCTGTGCAGCGGGGCGGGGCCCCAGCC 202
                ||| |||||| ||| :::
Db   173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-------GlySerValSer 189

Qy   203 AAGCAGCAGCAAGTGACCATCCCCTTCGGCTCTGGTGCCCCACCCAGTATGTGGGTGCC 262
                ||||| |||| |||:::||:::||||| |||||::||
Db   190 LysGlnLysProCysAspGlyLeuProLeuArgLeuLeuValProThrGlnPheGlnValGlyAla 209

Qy   263 ATTATTGCAAGGAGGGGGCCACCATCCACAACATCACAAAACAGACCCAGTCCAAGATA 322
                ||| ||||||::: |||||||::::||||:::
Db   210 IleIleGlyLysGlyGluGlyGluGlyLysAlaThrIleArgAsnIleThrLysGlnSerLysSerIle 229

Qy   323 GACGTGCATAGGAAGGAAGCAACCAGGTTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC 382
                ||||| |:::||||::|||| |||||:::||:::|||:::::||
Db   230 AspValHisArgLysGluAlaThrArgLeuGlnLeuLysLysProSerValCysThrProThr 249

Qy   383 CCTGAGGGCTGCTCCTCCGCTTGTAAGATGACTTGAGGAATTATGCATAAAGAGGCTAAG 442
                ||||||||| | |||||| ||| || :::||||
Db   250 ProGluGlyThrSerAlaAlaCysLysSerLeuLeuValSerIleLeuGluIleMetHisLysGluAlaGln 269

Qy   443 GACACCAAAAACGGCTGACGAGGTTCCCCTGAAGATCCTGCCCCATATAACTTTGTAGGG 502
                   ||| || ||| | ||| :::
Db   270 AspIleLysPheThrGluLeuProLeuGlnThrValLysAsnLeuLysLysIleGluGlnAspThrPheAspAlaHisAlaHisAsnAsnPheValGly 289

Qy   503 CGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAA 562
                || ||||||| |||||||||
Db   290 ArgLeuIleGlyLysGluGlyGlyArgAsnGlyArgAsnLeuLysLysLysIleGluGlnAspThrPheAspThrLys 309

Qy   563 ATCACCAACTCCCGTTGCAAGACCTTCACCCTTACACCCTGTG 622
                | ||| |||
Db   323 IleThrIleSerProLeuGlnGlnLeuThrLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

Db   310 IleThrIleSerProLeuGlnGlnLeuThrLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

Qy   623 AAGGGGGCCATCAGAGAATTGTGCAGGGCCGAGCCAGGAAATAATGAGAAAGTTCGGGAG 682
                |||||| |||:::|| ||
Db   330 LysGlyLysAsnValGluThrCysAlaLysAlaGluGlyGluIleMetLysLysLysIleArgGlu 349

Qy   683 GCCTATGAGAATGATGTGGCTGCCATGAG---------TCTCACCTGATCCTGGCCTGAAC 736
                :::||||||||||||||||||| ||||||
Db   350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 369

Qy   737 CTGGTGCTGTAGGTCTTTTCCAGGTTCATCCAGCCCCGCG-------CCT 787
                ||||||||||| |||||| :::||||||
Db   370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyPro 389

Qy   788 CCCAGCAGGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAG 847
                |||||||||||:::||| || :::|||||
Db   390 ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu 406

Qy   848 ATGGTGCAGGTGTTTATCCCCGCCCAGGCAGTGGCGCCATCATCGGCAAGAAGGGGCAG 907
                ||||||||||| ||| |||||| ||||||
Db   407 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnIleGlnGln 426

Qy   908 CACATCAAAACAGCTCTCCCCGGTTGCAGCAGCCCTCATCAAGATTCCATCAAGAATCCGAAACA 967
                ||||||||||| ||| ||||||
Db   427 HisIleLysGlnLeuSerArgPheAlaLeuGlyAlaSerIleLeuLysIleLeuAlaProAlaGluAla 446

Qy   968 CCTGACTCCAAAGTTCGTATGGTTATCATCAACTCAAGGAGGAGGAGAACTTCTTTGGTC 1027
                ||||| || ||| ||||||
Db   447 ProAspAlaLysValArgMetValIleIleIleThrGlyProProGluAlaGlnPheLysAla 466

Qy   1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGGAGAACTTCTTTGGTCCCAAGGAAGTG 1087
                ||||||||||| |||||||||||||||
Db   467 GlnGlyArgIleTyrGlyLysLysIleLysGluGluGluAsnPheValSerProLysLysGluValVal 486

Qy   1088 AAGCTGGAGCACCACATACGTGTGCCAGCATCAGCAGCTGCCGGGTCATTGGCAAAGGT 1147
                ||||||||| :::
Db   487 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyValLysLysGly 506

Qy   1148 GGAAAAACGGTGAACGAGTTGCAGAATTGACGGGACGTGAGGTGTGTAGTACCAAGAGAC 1207
                ||||||||| ||| ||||||||| |||
Db   507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProValValProArgAsp 526

Qy   1208 CAGACCCCTGATGAGACGACCAGGTCATCGTGAAAATCATCGGACATTCTATGCCAGT 1267
                |||||||||||||||||||| ||||
Db   527 GlnThrProAspGluAsnAspGlnValValLysIleIleThrGlyIleHisProProPheTyrAlaCys 546

Qy   1268 CAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAG---CAGCATCAG 1324
                |||||||||:::|||||||||| |||:::||
Db   547 GlnValAlaAlaGlnArgLysIleArgLysIleLeuGlnIleLeuGluThrGlnValLysGlnHisGlnGln 566

Qy   1325 AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGGAAG 1360
                ||::::
Db   567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579

RESULT 12
ABB75048
ID   ABB75048 standard; protein; 586 AA.
XX
AC   ABB75048;
XX
DT   01-MAY-2002 (first entry)
XX
DE   Human lung tumour L523S recombinant protein sequence SEQ ID NO:427.
XX
KW   Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW   immune response.
XX
OS   Homo sapiens.
XX
PN   WO200200174-A2.
XX
PD   03-JAN-2002.

```
XX
PF   28-JUN-2001; 2001WO-US021065.
XX
PR   28-JUN-2000; 2000US-00606421.
PR   02-AUG-2000; 2000US-00630940.
PR   21-AUG-2000; 2000US-00643597.
PR   15-SEP-2000; 2000US-00662786.
PR   09-OCT-2000; 2000US-00685696.
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,   Skeiky YAW,  Li SX,  Kalos MD,    Henderson RA;
PI   Mcneill PD, Fanger N,  Retter MW, Marnerakis M, Fanger GR;
PI   Vedvick TS, Carter D,  Watanabe Y,  Peckham DW;
XX
DR   WPI; 2002-090513/12.
DR   N-PSDB; ABL49283.
XX
PT   Polynucleotides encoding lung tumor polypeptides, useful for treating
PT   lung cancer or stimulating an immune response.
XX
PS   Claim 2; Page 354-355; 374pp; English.
XX
CC   The present invention describes human lung tumour proteins. Human lung
CC   tumour proteins and polynucleotides have cytostatic and immunostimulant
CC   activities, and can be used in vaccine production. Compositions
CC   comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC   proteins, T cell populations, or antigen presenting cells that express
CC   the lung tumour proteins are useful for treating lung cancer or
CC   stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC   ABB75070 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 586 AA;

Alignment Scores:
Pred. No.:            2.38e-144      Length:          586
Score:               1639.00        Matches:         336
Percent Similarity:    83.89%       Conservative:     44
Best Local Similarity: 74.17%       Mismatches:       57
Query Match:           52.70%       Indels:           16
DB:                     5            Gaps:              9
```

```
US-09-270-437D-5 (1-1708) x ABB75048 (1-586)

Qy     32 CGGGGAGCCATCATGAAGCTGAAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC  91
           |||| ||:::   |||   |||||||||||||||||||||||||||||||  ||||||||||
Db    140 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 159

Qy     92 TACATCCCGATGAGCAGATAGCA--CAGGGACCTGAGAATGGGCGCCAGG-----  142
           ||||||||||| |||| ||||||  |||||||||||||||||||||||||
Db    160 TyrIleProAspGluThrAlaAlaGlnGlyProLeuGlnGlnProArgGlyArgArg 179

Qy    143 GGCTTTGGCTCTCGGGGTCAGCCCGCCAGGGGCTCACCTGTGGCAGCGGGGGCCCCAGCC 202
           |||||||||||||||||||||:::   |||||||||||||||:::  |||::
Db    180 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro--------GlySerValSer 196

Qy    203 AAGCAGCAGCAAGTGGCAGATCACATCCCCCTTCGGCTTCCTTCGTGTGCCCACCCCGTGGC 262
           |||||::  |||||||||||||||||||||||:::|:||||||  |||||||||||
Db    197 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrProPheValGlyAla 216

Qy    263 ATTATTGGCAAGGAGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATA 322
           |||||||||||||||||| ||||||||||||||||:::||||| |||||||  ||||
Db    217 IleIleGlyLysGluGluGlyGlyAlaThrIleArgAsnIleThrLysGlnSerlysIle 236

Qy    323 GACGTGCATAGGAAGGAGAACGCAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC 382
           |||||||||||||||||||||||:::   ||||:::|||||||||| ||||||||:::
Db    237 AspValHisArgLysGluAsnAlaAlaGlyAlaGluAlaGlyLysSerIleThrIleLeuSerThr 256

Qy    383 CCTGAGGGCTGCTCCTCCGCTTGTGAAGATCTTGGAGATTATGCATAAAAGAGAGCTAAG 442
           ||||||||||||||||||||||||||||||||||||||||||||||  :::
```

```
Db   257 ProGluGlyThrSerAlaAlaCysLysLysSerIleLeuGluIleMetHisLysGluAlaGln 276

Qy   443 GACACCAAAACGGCTGACGAGGTTCCCTGAAGATCCTGCCCATAATAACTTTGTAGGG 502
          :::||||||||||||||||||||||||||||||:::||||
Db   277 AspIleLysPhePheThrGlnGluIleProLeuProLeuLysIleLeuAlaHisAsnAsnPheValGly 296

Qy   503 CGTTTCATTGGCCAAGGAGGACGGAACCTGAAGAAGTTAGAGCAAGATACCGACACAAA 562
          |||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db   297 ArgLeuIleGlyLysGluLysGluGlyIleGluGlnLysLysLysIleGluGlnThrAspThrLys 316

Qy   563 ATCACCATCTCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGACGACCATCACTCTG 622
          |||||||||||||  |||||||||||:::|||||||||||||||||:::
Db   317 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 336

Qy   623 AAGGGGGCCATCGAGAATGTTGCGAGGGCCGAGCAGGAAAATAATGAGAAAGTTCGGCAG 682
          ||||||  :::|||   ::||||||||||||||||||||||||||||||
Db   337 LysGlyAlaAsnValGluThrCysAlaLysAlaGluGlnGluLysAsnLysIleLysArgGlu 356

Qy   683 GCCTATGAGAATGATGTGGCTGCCATGAGC-------TCTCACCTGATCCCTGGCCTGAAC 736
          :::|||||||||||||||||||||||||||       ||||||||||||
Db   357 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn 376

Qy   737 CTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCAGTCCCGCG--------CCT 787
          |||||||||||||   :::||||||||||||||||||
Db   377 LeuAlaAsnAlaLeuGlyLeuPheLeuProThrSerGlyMetProProProSerGlyPro 396

Qy   788 CCCAGCAGCGTACTGGGGCGTGCTGCCCTATAGCTCCTTTATGCAGGGTCCCCAGGAGCAGGAG 847
          ||||||||||  ||||||||||||:::||
Db   397 ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer--GluThrGlu 413

Qy   848 ATGGTGCAGGTGTTATCCCGGCCAGGCAGTGGGCGGCCATCATCGGCAAGAGGGGCCAG 907
          :::||||||||||||  |||||||||||||||||||||||||||||||
Db   414 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnGln 433

Qy   908 CACATCAAAACAGCTCTCCCGGTTGCACCGCGCCTCCATCAAGAATTGCACACCCGAAACA 967
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   434 HisIleLysLysGlnLeuSerArgPheAlaGlyAlaPheLysSerIleLysIleAlaProAlaGluAla 453

Qy   968 CCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCGCGCAGAGGCCAATTCAAGGCT 1027
          ||||||||||||||||||||||||||||||||||||||:::|||||||||
Db   454 ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla 473

Qy   1028 CAGGGAAGAAGAATCTATGGCAAACTCAAGGAGGAGACTCTTCGTGTCCCAAGAGGAAGTG 1087
          ||||||||||||||:::||||||||||||||||||||||||||||||||
Db   474 GlnGlyArgArgIleTyrGlyLysLeuLysGluGluAsnPheValSerProLysGluVal 493

Qy   1088 AAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCGCGGGTCATTGGCAAGGT 1147
          ||||||||||||||:::   ||||||||||||||:::
Db   494 LysLeuGluAlaHisIleArgValProSerPheAlaGlyAlaGlyArgValIleGlyLysGly 513

Qy   1148 GGAAAAACCGTGAACGAGTTGCAGAATTTGACGGCAGTGAGGTGGTAGTACCAAGAGAC 1207
          |||||||||||||||||||||||||||||||||||||||||||||
Db   514 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgArgAsp 533

Qy   1208 CAGACCCCTGATGAGAACAGCAGGTCATCGTGAAAATCATCGGACATTTCATGCCCAGT 1267
          |||||||||||||||||  |||||:::|||||||||||||||||||
Db   534 GlnThrProAspGluAsnAspGlnValValLysIleIleGlyHisSerPheTyrAlaCys 553

Qy   1268 CAGATGGCTCAACGGAAGATCCGAGACATCCTGCCCAGGTTAAGCAG---CAGCATCAG 1324
          |||||||||||||||||||||||||||||:::||||  ||||||||||
Db   554 GlnValAlaGlnArgLysIleLysGluLeuLeuThrGlnValValLysGlnHisGlnGln 573

Qy   1325 AAGGGA--CAGAGTAACCAGGGCCCAGGCACGGAGGAAG 1360
          ||||||  ||||||||||||||||:::||||
Db   574 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 586
```

```
RESULT 13
ABP61968
ID   ABP61968 standard; protein; 586 AA.
XX
AC   ABP61968;
XX
DT   07-OCT-2002  (first entry)
```

XX
DE   Human lung cancer associated protein sequence SEQ ID NO:427.
XX
KW   Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200247534-A2.
XX
PD   20-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-US047576.
XX
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
PR   28-JUN-2001; 2001US-00897778.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI   Mcneill PD, Fanger MW, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI   Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR   WPI; 2002-583465/62.
DR   N-PSDB; ABQ92469.
XX
PT   Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT   the polynucleotides, useful in pharmaceutical compositions such as
PT   vaccines and as markers to indicate the presence of lung cancer.
XX
PS   Claim 9; Page 361-362; 381pp; English.
XX
CC   The present invention describes isolated human lung carcinoma
CC   polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC   activity, and can be used in gene therapy and in vaccines. Compositions
CC   comprising (I) or (II) can be used for stimulating an immune response in
CC   a patient and for treating lung cancer in a patient. Oligonucleotides of
CC   (I) can be used for detecting the presence of a cancer in a patient, by
CC   obtaining a biological sample from the patient, contacting the biological
CC   sample with the oligonucleotide, detecting in the sample, an amount of
CC   polynucleotide that hybridises to the oligonucleotide and comparing the
CC   amount of polynucleotide that hybridises to the oligonucleotide to a
CC   predetermined cut-off value, and determining the presence of a cancer in
CC   the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC   vaccines. (I) is useful as a marker to indicate the presence or absence
CC   of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC   ABP61992 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 586 AA;

Alignment Scores:
Pred. No.:   2.38e-144
Score: 1639.00
Percent Similarity: 83.89%
Best Local Similarity: 74.17%
Query Match: 52.70%
DB:

US-09-270-437D-5 (1-1708) x ABP61968 (1-586)

Length: 586
Matches: 336
Conservative: 44
Mismatches: 57
Indels: 16
Gaps: 9

US-09-270-437D-5 (1-1708) x ABP61968 (1-586)

```
Qy      32  CGGGGAGGCCATCATGAAGCTGAATGGGCACCAGTTGGAGAACCATGCCTGAGAAGGTCTCC      91
            ||||||| |:::   |||||||||||   ||||| ||     |||||||||||||||
Db     140  ArgGlnAlaLeuAspLysLeuAsnGlyHisGlnLeuGluLeuAsnPheThrLeuLysValAla     159

Qy      92  TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCCCGAGGG------    142
            |   |||  |||||  ||   |||      |||
Db     160  TyrIleProAspGluThrAlaAlaGlnGlnAsnProLeuGlnIleProGlyArgGlyArgArg     179

Qy     143  GGCTTTGGCTCTCGGGGGTCAGCCCCGCCAGGGCTCACCTGTGCAGCGGGGCCCCAGGCC    202
            |  ||| ||||   |||      |||     ::: ||||
Db     180  GlyLeuGlyGlnArgGlySerArgGlnGlySerPro-------GlySerValSer     196
```

```
Db    554 GlnValAlaGlnLysIleArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGln 573
          |||::||||||||||||:::::||||  |||||||||  |||||||||      |||

Qy   1325 AAGGGA--CAGAGTAACCAGGCCCAGGCACGGAGGAAG 1360
          ||:::|||| ||:|||||||||||
Db    574 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 586
```

RESULT 14
ADA28517
ID   ADA28517 standard; protein; 586 AA.
XX
AC   ADA28517;
XX
DT   20-NOV-2003　(first entry)
XX
DE   Recombinant human lung tumour protein L523S.
XX
KW   cancer; lung cancer; gene therapy; vaccine; human;
KW   lung squamous cell carcinoma.
XX
OS   Homo sapiens.
XX
PN   US2003064947-A1.
XX
PD   03-APR-2003.
XX
PF   30-NOV-2001; 2001US-00007700.
XX
PR   18-MAR-1998;　98US-00040802.
PR   27-JUL-1998;　98US-00123912.
PR   22-DEC-1998;　98US-00221107.
PR   02-APR-1999;　99US-00285479.
PR   17-DEC-1999;　99US-00466396.
PR   30-DEC-1999;　99US-00476496.
PR   10-JAN-2000;　2000US-00480884.
PR   22-FEB-2000;　2000US-00510376.
PR   04-APR-2000;　2000US-00542615.
PR   28-JUN-2000;　2000US-00606421.
PR   02-AUG-2000;　2000US-00630940.
PR   21-AUG-2000;　2000US-00643597.
PR   15-SEP-2000;　2000US-00662786.
PR   09-OCT-2000;　2000US-00685696.
PR   12-DEC-2000;　2000US-00735705.
PR   07-MAY-2001;　2001US-00850716.
PR   28-JUN-2001;　2001US-00897778.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,　Wang A,　Skeiky YAW,　Li SX,　Kalos MD,　Henderson RA;
PI   Mcneill PD,　Fanger N,　Retter MW,　Durham M,　Fanger GR,　Vedvick TS;
PI   Carter D,　Watanabe Y,　Peckham DW,　Cai F,　Foy TM;
XX
DR   WPI; 2003-540798/51.
XX
PT   New isolated polynucleotides and polypeptides useful for diagnosing,
PT   preventing and/or treating cancer, particularly lung cancer.
XX
PS   Claim 9; Page 273-275; 296pp; English.
XX
CC   The invention describes isolated polynucleotides and polypeptides useful
CC   for diagnosing, preventing and/or treating cancer, particularly lung
CC   cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC   defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
CC   specification; complements of the nucleotide sequences cited above; at
CC   least 10 contiguous residues of the nucleotide sequences cited above; a
CC   sequence that hybridise to any of the nucleotide sequences under highly
CC   stringent conditions; a sequence that is at least 75 or 90% identical to
CC   the above nucleotide sequences; or degenerate variants of the above
CC   nucleotide sequences. The composition and methods are useful in
CC   diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC   in gene therapy and in vaccines. This is the amino acid sequence of a
CC   recombinant human lung tumour associated protein.
XX

SQ   Sequence 586 AA;

Alignment Scores:
Pred. No.:　　　　　　　　2.38e-144　　　Length:　　　　586
Score:　　　　　　　　　　1639.00　　　　Matches:　　　336
Percent Similarity:　　　83.89%　　　　Conservative:　44
Best Local Similarity:　74.17%　　　　Mismatches:　　57
Query Match:　　　　　　　52.70%　　　　Indels:　　　　16
DB:　　　　　　　　　　　 7　　　　　　　Gaps:　　　　　9

US-09-270-437D-5 (1-1708) x ADA28517 (1-586)

Qy   32 CGGGGAGCCATCATGAAGGTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC 91
         ||| ||::::　||||||||||||||||||||||||
Db   140 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 159

Qy   92 TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCCGAGGG------- 142
         ||||||||| |||||||||||| ||| ||
Db   160 TyrIleProAspGluThrAlaAlaGlnGlnAlaGlnProLeuGlnLeuAsnProArgGlyArgArg 179

Qy   143 GGCTTTGGCTCTCGGGGTCAGCCCCCAGGGTCACCTGTGGCAGCGGGGGCCCCAGCC 202
         ||||||||||||||||||||||||||||||||||||||| :::::　:
Db   180 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro---------GlySerValSer 196

Qy   203 AAGCAGCAGCAAGTGGACATCCCCCTTGGCTCCTGGTGCCCACCCAGTATGTGGGTGCC 262
         ||||||:::::　|||||||||||||||||||||||||||||||||||||||||
Db   197 LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAla 216

Qy   263 ATTATTGGCAAGGAGGAGGGCCACCATCCCAACATCCAAACAGACCCAGTCCAAGATA 322
         |||||||| ||||||||||||||||| ||||||| ||||||||||||||||||||||
Db   217 IleIleGlyLysGluGluGlyHisHisProAsnIleGlnThrLysGlnThrGlnSerLysIle 236

Qy   323 GACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC 382
         |||　|||||||||||||| ||||||||||||||||||||||||||||||||||
Db   237 AspValHisArgLysGluAsnAlaGlyAlaAlaGluLysAlaIleSerValHisSerThr 256

Qy   383 CCTGAGGGCTGCTCCTCCGGTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG 442
         |||||||||||| ::::　||||||||||||||||||||||| :::
Db   257 ProGluGlyThrSerAlaAlaCysLysLysSerIleLeuGluIleMetHisLysGluAlaGln 276

Qy   443 GACACCAAAACGGCTGACGAGGTTCCCGAGATCCTGGCCCATAATAACTTTGTAGGG 502
         |||||　:::||||||||||||||||||||||||||||||||||||||||||
Db   277 AspIleLysPheThrGluGluValProGluIleLeuAlaHisAsnAsnPheValGly 296

Qy   503 CGTCTCATTGGCAAGGAAGACGGAACCTGAGAAGGTAGAGCAAGATACCGAGACAAAA 562
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   297 ArgLeuIleGlyLysGluGluGlyGlyArgAsnLeuLysLysLeuGluGlnAspThrArgThrLys 316

Qy   563 ATCACCATCTCCTCGTTGCAAGACTTACCCTTACACACCCTGAGAGGACCATCACTCTG 622
         ||||||||||| :::||||||||||||||||||||||||||||||||||　------CCT 787
Db   317 IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 336

Qy   623 AAGGGGGCCATCAGAAATTGTTGCGAGGGGCCGAGCAGAAATATGAAGAAGTTCGGGAG 682
         ||||||||||||||||| :::||||||:::||||||||||||||:: :||
Db   337 LysGlyAsnValGluThrCysAlaLysAlaGluGluGluMetLysLysIleArgGlu 356

Qy   683 GCCCATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCTGCCTGAAC 736
         :::||||||||||||||||||||||||||　　::::|||||||||||||||||||||||
Db   357 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleLeuProGlyLeuAsn 376

Qy   737 CTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCG-------CCT 787
         ||| |||||||| :::|||||||| ::|||
Db   377 LeuAsnAlaLeuGlyLeuPheProThrSerGlyMetProProThrSerGlyPro 396

Qy   788 CCCAGCAGCGTTACTGGGGCTGCTCCTATAGCTCCTTATGCAGCTCCCGAGCAGGAG 847
         ||||||||||||||||||||||||||||||||||
Db   397 ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer--GluThrGlu 413

Qy   848 ATGGTGCAGGTGTTTATCCCGCCAGGCAGTGGGCCGCCATCATCGGCAAGAAGGGGCAG 907
         ||| :::|||||||||||| ::||||||||||:::|||
Db   414 ThrValHisLeuPheIleProAlaLeuSerValGlyAlaAlaIleIleGlyLysLysGlyGln 433

Qy   908 CACATCAAAACAGCTCTCCCGGTTTGCCAGGCGCCTCCATCAAGATTGCACCACCGGAAACA 967
```

CC   which have cytostatic activity. The polypeptides and polynucleotides are
CC   used in compositions and vaccines to inhibit the development of cancer,
CC   especially lung cancer, in a patient. Methods described in the invention
CC   can be used to monitor the progression of a cancer by carrying out the
CC   detection at subsequent time points and comparing the results from the
CC   different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC   are treated with P2, polynucleotides encoding P2 or antigen presenting
CC   cells expressing P2 and then administered to the patient to inhibit
XX   development of cancer
SQ   Sequence 579 AA;

Alignment Scores:
Pred. No.:          3.64e-144      Length:            579
Score:              1637.00        Matches:           336
Percent Similarity: 83.66%         Conservative:       43
Best Local Similarity: 74.17%      Mismatches:         58
Query Match:        52.64%         Indels:             16
DB:                                Gaps:                9

US-09-270-437D-5 (1-1708) x AAB11328 (1-579)

Qy     32 CGGGAGCCATCATGAAGCTGAATGCCACCAGTTGGAGAACCATGCCCGAAGGTCTCC   91
          |||||||| |||| |||   |||||||||||||||||||||           ||| :::
Db    133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheLeuLysValAla  152

Qy     92 TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGGCCAGGG----  142
          |||||||
Db    153 TyrIleProAspGluMetAlaAlaGlnAlaGlnAsnProLeuGlnProArgGlyArgArg  172

Qy    143 GGCTTTGGCTCTCGGGGTCAGCCCGCCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCC  202
          |||    :::
Db    173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro-------GlySerValSer  189

Qy    203 AAGCAGCAGCAAGTGGACATCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCC  262
          ||||:::
Db    190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla  209

Qy    263 ATTATTGGCAAGGAGGGGCCACCATCGCACATCACAAACAGACCCAGTCCAAGATA  322
          |||||
Db    210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrArgLysGlnThrGlnSerLysIle  229

Qy    323 GACGTGCATAGGAAGAGAAGAACGCAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC  382
          ||||:::
Db    230 AspValHisArgLysArgLysArgThrGlnCysSerLeuLysLysSerIleLeuThrLeuSerThr  249

Qy    383 CCTGAGGGCTGCTCCTCCCGGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAG  442
          ||||:::
Db    250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGln  269

Qy    443 GACACCAAAACGGCTGCGACGAGTTCCCTGAAGATCCTGGCCCATAATAACTTGTAGGG  502
          |||    :::
Db    270 AspIleLysPheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGly  289

Qy    503 CGTCTCATTGGCAAGGAAGCACGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAA  562
          ||||
Db    290 ArgLeuIleGlyLysGluGlyArgAsnLeuGlyLysLeuGluGluAspThrArgThrLys  309

Qy    563 ATCACCACCTTCCTCGTTCGTGGTCTGCAAGACCTTACCCTTTACAACCCTGAGGAGGACCATCACCGTG  622
          |||
Db    310 IleThrIleSerProLeuGlnLeuThrLeuTyrAsnProLeuArgThrIleThrVal  329

Qy    623 AAGGGGGCCATCGAGAATTGTTGCAGGGGCGAGCAGGAAATAATGAAGAAAGTTCGGGAG  682
          |||||    :::
Db    330 LysGlyLysAsnValGluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGlu  349

Qy    683 GCCTATGAGAATGATGTGGCCATGAGC-----TCTCACCTGATCCCTGGCCTGAAC  736
          :::
Db    350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProLeuGluLeuAsn  369

Qy    737 CTGGCTGCTGTGTAGGTCTTTTCCAGTTCATCCAGCGCAGTCCCGCCG-------CCT  787
          |||||||
Db    370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyPro  389

Db    434 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLeuAlaProAlaGluAla  453

Qy    968 CCTGACTCCAAGTTCGTATGGTTATCATCACCTGGACACCGCCAGAGGCCCAATTCAAGGCT  1027
          |||||:::||||||||||||||||
Db    454 ProAspAlaLysValArgMetValIleIleThrGlyThrProProGluAlaGlnPheLysAla  473

Qy   1028 CAGGGAGAATCTATGCCAAAATCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG  1087
          ||||||:::||||
Db    474 GlnGlyArgIleTyrGlyIleTyrGlyIleLysLysGluGluAsnPheValSerProLysGluGluVal  493

Qy   1088 AAGCTGGAGACCCACCATACGTGTGCCACAGTCAGCAGCTGGCCCGGGTCATTGGCAAAGGT  1147
          ||||||:::
Db    494 LysLeuGluAlaHisHisIleLeuArgValProSerPheAlaGlyArgValIleGlyLysGly  513

Qy   1148 GGAAAAACGGTGAACAGTTGCAGAATTTGACCGCAGCTGAGTGGTGGTAGTACCAAGAGAC  1207
          ||||:::
Db    514 GlyLysThrValAsnGluLeuGlnAsnLeuserSerAlaGluValValProArgArgAsp  533

Qy   1208 CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGAACATTCTATGCCAGT  1267
          ||||
Db    534 GlnThrProAspGluAsnAspGlnValIleValLysIleIleThrGlyHisPheTyrAlaCys  553

Qy   1268 CAGATGGCTCAACGGAGAATCCGAGACATCCTGGCCCCAGGTTAAGCAG---CAGCATCAG  1324
          ||||:::
Db    554 GlnValAlaGlnArgLysIleArgLysSerIleGluGlnIleLeuLeuThrGlnIleLysLeuGlnHisGln  573

Qy   1325 AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGAAG  1360
          |||    ||||
Db    574 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys  586

RESULT 15
AAB11328
ID   AAB11328 standard; protein; 579 AA.
XX
AC   AAB11328;
XX
DT   21-FEB-2001 (first entry)
XX
DE   Human lung cancer-associated protein L523S.
XX
KW   Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW   vaccine; detection.
XX
OS   Homo sapiens.
XX
PN   WO200061612-A2.
XX
PD   19-OCT-2000.
XX
PF   03-APR-2000; 2000WO-US008896.
XX
PR   02-APR-1999;   99US-00285479.
PR   17-DEC-1999;   99US-00466396.
PR   30-DEC-1999;   99US-00476496.
PR   10-JAN-2000;   2000US-00480884.
PR   22-FEB-2000;   2000US-00510376.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T,  Fan L;
XX
DR   WPI; 2000-628399/60.
DR   N-PSDB; AAC65900.
XX
PT   Isolated polypeptide comprising an immunogenic portion of a lung tumor
PT   protein is used for detecting and monitoring progression of lung cancer
PT   in a patient.
XX
PS   Claim 3; Page 186-188; 261pp; English.
XX
CC   This invention describes a novel isolated polypeptide (I) which
CC   comprising an immunogenic portion of a lung tumor protein or variant (P2)

Wed Jul 14 08:56:04 2004

us-09-270-437d-5_1.rag

```
Qy    788  CCCAGCAGCGGTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAG  847
                |||:|||||:::::||||||                   |||||  :::  |||||||
Db    390  ProSerAlaMetThr------ProProTyrProGlnPheGlnSer---GluThrGlu  406

Qy    848  ATGGTGCAGGTGTTTATCCCCCCCAGGCAGTGGGCCCATCATCCGCAAGAAGGGGCAG  907
                |||   ::|||||||||||||    ::|||||||||||||:::||||||||||
Db    407  ThrValHisGlnPheIleProAlaLeuSerValGlyAlaIleIleIleGlyGlnGlyGln  426

Qy    908  CACATCAAACAGCTCTCCCGGTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACA  967
                |||||||||||||||||||||||||||||||||||||||||||||||      |||
Db    427  HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysAlaProAlaGluAla  446

Qy    968  CCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGCCCCAATTCAAGGCT  1027
                ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db    447  ProAspAlaLysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAla  466

Qy    1028 CAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG  1087
                ||||||||||||||||||||:::||||||||||:|||||||||||||||||||||
Db    467  GlnGlyArgIleTyrGlyLysLeuLysGluGluAsnPheValSerProLysGluGluVal  486

Qy    1088 AAGCTGGAGACCCACATACGTGCCAGCAGCTGCCGGTCATTGGCAAAGGT  1147
                ||||||||||||||||||||||||||||||||:::||||||||||||||
Db    487  LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyValIleGlyLysLysGly  506

Qy    1148 GGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTGGTAGTACCAAGAGAC  1207
                ||||||||||||||||:::::|||||||||||:::|||||||||||||||||||
Db    507  GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProGlnArgAsp  526

Qy    1208 CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTCTATGCCAGT  1267
                ||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db    527  GlnThrProAspGluAsnAspGlnValValIleValLysIleIleThrGlyHisPheTyrAlaCys  546

Qy    1268 CAGATGGCTCAACGAGGAGATCCGAGACATCCTGGCCCAGGTTAAGCAG---CAGCATCAG  1324
                ||:::||||||||||:::||||||||||||||||||||:::|||||||   |||||||
Db    547  GlnValAlaGlnArgGlyIleLeuGlnIleLeuAlaGlnValLysGlnHisGlnGlnGln  566

Qy    1325 AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGGAAG  1360
                ||:::    |||:::||||||||||||||
Db    567  LysAlaLeuGlnSerGlyLysProProGlnSerArgArgLys  579
```

Search completed: July 13, 2004, 12:08:40
Job time : 108 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 13, 2004, 12:07:12 ; Search time 22 Seconds
                                   (without alignments)
                                   8016.105 Million cell updates/sec

Title: US-09-270-437D-5
Perfect score: 3110
Sequence: 1 aggacggctgccgcaccgcc.........atttcctcaggtttaaaa 1708

Scoring table: BLOSUM62
            Xgapop 10.0 , Xgapext   0.5
            Ygapop 10.0 , Ygapext   0.5
            Fgapop 6.0 , Fgapext   7.0
            Delop   6.0 , Delext   7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_13072004_121922_9641/app_query.fasta_1.1863
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN_1.1.33 @runat_13072004_121922_9641 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   Issued_Patents_AA:*
       1: /cgn2_6/ptcdata/2/iaa/5A_COMB.pep:*
       2: /cgn2_6/ptcdata/2/iaa/5B_COMB.pep:*
       3: /cgn2_6/ptcdata/2/iaa/6A_COMB.pep:*
       4: /cgn2_6/ptcdata/2/iaa/6B_COMB.pep:*
       5: /cgn2_6/ptcdata/2/iaa/PCTUS_COMB.pep:*
       6: /cgn2_6/ptcdata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2208	71.0	577	3	US-09-261-855-2	Sequence 2, Appli
2	1639	52.7	579	4	US-09-643-597-348	Sequence 348, App
3	1639	52.7	579	4	US-09-542-615A-348	Sequence 348, App
4	1639	52.7	579	4	US-09-606-421B-348	Sequence 348, App
5	1637	52.6	579	4	US-09-643-597-176	Sequence 176, App
6	1637	52.6	579	4	US-09-480-884A-176	Sequence 176, App
7	1637	52.6	579	4	US-09-542-615A-176	Sequence 176, App
8	1637	52.6	579	4	US-09-606-421B-176	Sequence 176, App
9	248.5	8.0	643	1	US-08-021-608D-10	Sequence 10, Appl
10	248.5	8.0	643	5	US-08-726-160-10	Sequence 10, Appl
11	248.5	8.0	643	5	PCT-US94-01782-10	Sequence 10, Appl
12	245	7.9	49	3	US-09-261-855-18	Sequence 18, Appl
13	245	7.9	245	1	US-08-021-608D-2	Sequence 2, Appli
14	245	7.9	245	1	US-08-726-160-2	Sequence 2, Appli
15	245	7.8	245	5	PCT-US94-01782-2	Sequence 2, Appli
16	242	7.8	242	1	US-08-021-608D-8	Sequence 8, Appli
17	242	7.8	242	5	US-08-726-160-8	Sequence 8, Appli
18	242	7.6	236	5	PCT-US94-01782-8	Sequence 8, Appli
19	236	7.4	229	3	US-09-261-855-20	Sequence 20, Appl
20	229	7.3	228	3	US-09-261-855-17	Sequence 17, Appl
21	228	7.1	222	3	US-09-261-855-19	Sequence 19, Appl
22	222	7.1	221.5	3	US-09-261-855-22	Sequence 22, Appl
23	221.5	7.1	219	1	US-08-187-793-4	Sequence 4, Appli
24	219	7.0	214	3	US-09-261-855-24	Sequence 24, Appl
25	214	6.9	191	3	US-09-261-855-21	Sequence 21, Appl
26	191	6.1	176.5	3	US-09-261-855-23	Sequence 23, Appl
27	176.5	5.7	174	4	US-09-252-991A-32367	Sequence 32367, A
28	174	5.6	162.5	1	US-08-187-793-2	Sequence 2, Appli
29	162.5	5.2	157	4	US-09-252-991A-24009	Sequence 24009, A
30	157	5.0	154.5	4	US-09-252-991A-30464	Sequence 30464, A
31	154.5	5.0	154.5	1	US-08-021-608D-6	Sequence 6, Appli
32	154.5	5.0	154.5	1	US-08-726-160-6	Sequence 6, Appli
33	154.5	5.0	154	5	PCT-US94-01782-6	Sequence 6, Appli
34	154	5.0	151.5	4	US-09-252-991A-19178	Sequence 19178, A
35	151.5	4.9	151.5	4	US-09-252-991A-19034	Sequence 19034, A
36	151.5	4.9	150	4	US-09-252-991A-16940	Sequence 16940, A
37	150	4.8	145	4	US-09-252-991A-28458	Sequence 28458, A
38	145	4.8	144.5	4	US-09-976-594-615	Sequence 615, App
39	144.5	4.7	143.5	4	US-09-252-991A-20408	Sequence 20408, A
40	144.5	4.6	142.5	4	US-09-252-991A-17231	Sequence 17231, A
41	143.5	4.6	141.5	4	US-09-252-991A-31974	Sequence 31974, A
42	142.5	4.6	141	4	US-09-252-991A-16789	Sequence 16789, A
43	141.5	4.5	141	4	US-09-252-991A-18035	Sequence 18035, A
44	141	4.5	141	2	US-08-935-450-6	Sequence 6, Appli
45	141	4.5		4	US-08-952-899-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-261-855-2
; Sequence 2, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-2

Alignment Scores:
Pred. No.:   8.33e-200   Length: 577
Score: 2208.00   Matches: 438
Percent Similarity: 98.65%   Conservative: 1
Best Local Similarity: 98.43%   Mismatches: 4
Query Match: 71.00%   Indels: 2
DB: 3   Gaps: 1

US-09-270-437D-5 (1-1708) x US-09-261-855-2 (1-577)

Qy   32 CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTCGAGAACCATGCCCTGAAGGTCTCC 91
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   133 ArgGlnAlaIleMetLysLeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSer 152

Qy   92 TACATCCCGATGAGCAGATAGCACAGGGACCTGAGAATGGCGCCGCCGAGGGGCTTGGC 151

Db   153 TyrIleProAspGluGlnIleThrGlnGlyProGluAsnGlyArgArgGlyPheGly 172

Qy 1226 GACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCGATGGCTCAACGGAAG 1285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 533 AspGlnValIleValLysIleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLys 552

Qy 1286 ATCCCAGACATCCTCGCCCAGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACCAGGCC 1345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 553 IleArgAspIleLeuAlaGlnLeuSerSerSerIleArgArgAspArgValThrArgPro 572

Qy 1346 CAGGCACGGAGGAAG 1360
          ||||||||||||
Db 573 GlnAlaArgArgLys 577

RESULT 2
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 626072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348

Alignment Scores:
Pred. No.:   Length: 579
Score: 4.72e-146   Matches: 336
Percent Similarity: 1639.00   Conservative: 44
Best Local Similarity: 83.89%   Mismatches: 57
Query Match: 74.17%   Indels: 16
DB: 52.70%   Gaps: 9
                    4

US-09-270-437D-5 (1-1708) x US-09-643-597-348 (1-579)

Qy   32 CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC 91
          ||| ::.   :|:::|:
Db 133 ArgGlnAlaLeuAspLysLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152

Qy   92 TACATCCCGATGAGCAGCAGTAGCA---CAGGGACCTGAGAATGGGCGCCGAGGG----- 142
          |||||| ||||||
Db 153 TyrIleProAspGluThrAlaAlaLeuAlaGlnGlnAsnProLeuGlnProArgGlyArgArg 172

Qy 143 GGCTTTGGCTCTGGGGTCAGGGTCAGCCCGCCAGGGGTCACCTGTGGCAGGCCCCCAGCC 202
          |||||| :::
Db 173 GlyLeuGlyGlyArgGlySerSerArgGlnGlySerPro------GlySerValSer 189

Qy 203 AAGCAGCAGCCATCATGAAGCTGGACATCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCC 262
          ||||||:.   ::::|
Db 190 LysGlnLysProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAla 209

Qy 263 ATTATTGGCAAGGAGGGGGCCAACATCCCAACTCACAAAACAGACCCAGTCAAGATA 322
          ||||||
Db 210 IleIleGlyLysGluGluGlyAlaThrIleArgAsnIleArgLysThrGlnThrGlnSerLysIle 229

Qy 323 GACGTGCATAGGAAGGAGAACCCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC 382
          ||:::|:::.
Db ...

; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-348

Alignment Scores:
Pred. No.:
Score: 1639.00   Length: 579
Percent Similarity: 83.89%   Matches: 336
Best Local Similarity: 74.17%   Conservative: 44
Query Match: 52.70%   Mismatches: 57
DB: 4   Indels: 16
                                              Gaps: 9

US-09-270-437D-5 (1-1708) x US-09-542-615A-348 (1-579)

Qy   32 CGGGGAGCCATCATGAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC   91
          ||||||||:||: ||||||
Db   133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla 152

Qy   92 TACATCCCCGATGAGCAGATAGCA----CAGGGACCTGAGAATGGGCGCCGAGG------- 142
          |||| ||| ||
Db   153 TyrIleProAspGluThrAlaAlaGlnIleGlnAlaGlnProArgGlyArgArg   172

Qy   143 GGCTTTGGCTCTCGGGGTCAGCCCCCCCCAGGGCTCACCTGTGGCAGGCGGGGCCCCAGCC 202
          |||| ||:::
Db   173 GlyLeuGlyLeuArgGlySerArgGlnGlySerPro--------GlySerValSer   189

Qy   203 AAGCAGCAGCAAGTGACATCCCCCTTCGGCTCCTTGTGTCCCAGTATGTGGCTGCC 262
          ||||||::: ||:::
Db   190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyGlyAla 209

Qy   263 ATTATTGGCAAGGAGGGGCCACCATCCGCAACATCCAGAAACGCAAGTCCAAGATA 322
          |||| |||
Db   210 IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysGlnSerLysLysIle 229

Qy   323 GACGTGCATAGGAAGGAGAACGCAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC 382
          |||| ||:::
Db   230 AspValHisArgLysGluAsnAlaAlaGlnLeuLysLysProSerIleLeuSerThr 249

Qy   383 CCTGAGGGCTGTCCTCCGGTTGTAAGATGCGTTGGAGATTATGCATAAAGAGGCTAAG 442
          ||||||| ||:::
Db   250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuMetHisLysGluLysAlaGln 269

Qy   443 GACACCAAAACGGCTGACGAGGTTCCCTGCAGATCCTGGCCCATAATAACTTTGTAGGG 502
          |||||| ||:::
Db   270 AspIleLysPheThrGluGluLeuProLeuGlyIleLeuAlaHisAsnAsnPheValGly 289

Qy   503 CGTCTCATTGGCAAGGAAGGAGGAACCCTGAAGAGAGGTAGAGCAAGATACCGAGACAAAA 562
          |||| ||||
Db   290 ArgLeuIleGlyLysGluGlyGlyThrLeuLysArgGlyArgAlaArgTyrArgAspLys 309

Qy   563 ATCACCAATCTCCTTGTTGCAAGACCTTACCCTTTACAAACCCTGAGAGACCATCACTGTG 622
          ||::: ||||
Db   310 IleThrThrLeuSerProLeuGlnThrLeuTyrAsnProGluArgThrIleThrVal 329

Qy   623 AAGGGGAGCCATCGAGAATTGTTGCAGGGCCGAGCAGAAATAATGAAGAAGTTCGGGAG 682
          |||||| ::||||

```
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-606-421B-348

Alignment Scores:
Pred. No.:                        4.72e-146
Score:                            1639.00
Percent Similarity:               83.89%
Best Local Similarity:            74.17%
Query Match:                      52.70%
DB:                               4

                    Length:       579
                    Matches:      336
                    Conservative: 44
                    Mismatches:   57
                    Indels:       16
                    Gaps:         9

US-09-270-437D-5 (1-1708) x US-09-606-421B-348 (1-579)

Qy     32  CGGGGAGCCATCATGAAGCTGATGACCCAGTTGGAGAACCATGCCCTGAAGGTCTCC   91
               :::||||||||||||||||||||||||  |||||||||  ||||||||||||
Db    133  ArgGlnAlaLeuAspLysProLeuLysAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla  152

Qy     92  TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCCGCCGAGG------   142
               ||||||:||||||:  :|||||:  |||||||||||||||||||||||
Db    153  TyrIleProAspGluThrArgAlaAlaGlnGlnAsnProLeuGlnAsnProArgGlyArgArg   172

Qy    143  GGCTTTGGCTCTGGGGTGCAGCCTCAGCGGCTCACCTGTGCGCAGGGCGCCCCAGCC  202
               ||||||||:|||||||||||||||||||:|||||||   ::::||:
Db    173  GlyLeuGlyGlnArgGlySerSerArgGlyGlySerPro--------GlySerValSer  189

Qy    203  AAGCAGCAGCAAGTGGACATCCCCTTCGGTCCCGGTGCCTCCCGGTATGTGGGTGCC   262
               |||||||||||:|||||||||||||||||:|::||||||||||||||||||||:
Db    190  LysGlnLysProCysAspLeuProLeuPheLeuArgGlnLysLeuArgProLeuValProVal  209

Qy    263  ATTATTGGCAAGAGAGGGGCCACCATCAGCGCAGTCCAACATCACAAAACAGAGCCAAGTCCAAGATA  322
               :|||||||||||:||||||||||||||||||||:::||||||||||||||||||||||
Db    210  IleIleGlyLysGlyGluGlyAlaThrIleArgAsnIleThrLysGlnThrGluSerLysIle  229

Qy    323  GACTGCATAGGAAGGAGAGGAGCAGGTGCAGCTGAGGTGGTGAGGTGCACTCCACC   382
               ||||:|||||||:|||||||||||:|||||||||||||::::||:|||||||||
Db    230  AspValHisArgArgArgArgArgSerAlaGlyAlaGlyAlaAlaGluLysLysSerIleThrIleLeuSerThr  249

Qy    383  CCTGAGAGGCTGCTCCTCCGCTTGTGAAGATCATCGGACATTTCTATGCCAGT   442
               |||||||||||||||||||:|||:|||||||||||||||||||::::||||:
Db    250  ProGluGlyThrSerAlaAlaCysLysSerIleLeuGluIleIleMetHisGluAlaGln  269

Qy    443  GACACCAAAACGGCTGAAGATCCGAAGATCTGGCCCATAATAACTTTGTAGGG   502
               |||||||||||:|||||||||||||||||||:|||||||||||||||||||
Db    270  AspIleLysPheThrGluGluIleProLeuLysIleLeuLeuAlaHisAsnAsnPheValGly  289

Qy    503  CGTCTCATTGGCAAGGAAGGACGAACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTG   562
               :|||||:||:::||||||||||||||||||||||||||
Db    290  ArgLeuIleGlyLysGluGlyValGlyArgAsnLeuLysGlyGlnAspThrAspThrLys  309

Qy    563  ATCACCATCTCCCGTTGCAAGACTTCAGAAGCTTTACCGTTGCAAGGGCTGAGAGAGACCATCACTGTG   622
               |||||||||||||||||:||::::||||||:||||||:::|||||||||||
Db    310  IleThrIleSerProLeuGlyLeuThrLeuTyrAsnProGluArgThrIleThrVal  329

Qy    623  AAGGGGCCCATCAGAATTGTTGCCAGGGGCCGAGCAGGAAATAATGAAGAAGTTCGGGAG   682
               |||||||||||||||:::||||||||||||||||||||||||||||||||
Db    330  LysGlyAsnValGluThrCysAlaLysAlaGluGluMetLysLysIleArgGlu  349

Qy    683  GCCTATGAGAATGATGTGGCTGCCATGAGC-------TCTCACCTGATCCCTGGGCCTGAAC   736
               :::||||||||||||||||||:::|||:   |||:::|||||||
Db    350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn  369

Qy    737  CTGGCTGCTGTGTAGGTCTTTTCCAGTTCATCCAGCGCAGTCCCGCG--------CCT   787
               |||||||||||||:::::||||||||||||||:::  |:||:
Db    370  LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro  389

Qy    788  CCCAGCAGGTTACTGGGCTGCTCCCTATAGCTCCTTATGCAGGGCTCCCAGCAGGAG   847
               |||||||::||||||   ||||:::  |||||||||||||||||||
Db    390  ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGluGlu  406

Qy    848  ATGGTGCAGGTGTTTATCCCGCCAGGCAGCAGTGGGCCGCCATCATCGGCAAGAAGGGGCAG   907
               ||||||||||||:::||||||||||||||||||||||||||||||||:::
Db    407  ThrValHisLeuPheLeuProAlaLeuSerValGlyAlaIleIleGlyLysLysGlyGln   426
```

```
Db    330  LysGlyAsnValGluThrCysAlaLysAlaGluGluMetLysLysIleArgGlu  349

Qy    683  GCCTATGAGAATGATGTGGCTGCCATGAGC-------TCTCACCTGATCCCTGGGCCTGAAC   736
               :::||||||||||||||||||:::|||:   |||:::|||||||
Db    350  SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsn  369

Qy    737  CTGGCTGCTGTGTAGGTCTTTTCCAGTTCATCCAGCGCAGTCCCGCG--------CCT   787
               |||||||||||||:::::||||||||||||||:::  |:||:
Db    370  LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProThrSerGlyPro  389

Qy    788  CCCAGCAGGTTACTGGGCTGCTCCCTATAGCTCCTTATGCAGGGCTCCCAGCAGGAG   847
               |||||||::||||||   ||||:::  |||||||||||||||||||
Db    390  ProSerAlaMetThr-----ProProTyrProGlnPheGluGlnSer---GluThrGluGlu  406

Qy    848  ATGGTGCAGGTGTTTATCCCGCCAGGCAGCAGTGGGCCGCCATCATCGGCAAGAAGGGGCAG   907
               ||||||||||||:::||||||||||||||||||||||||||||||||:::
Db    407  ThrValHisLeuPheLeuProAlaLeuSerValGlyAlaIleIleGlyLysLysGlyGln  426

Qy    908  CACATCAAAACAGCTCTCCCGGTTGCCAGCGCCTCCATCAAGATGCACCACCCGAAACA   967
               ||||||||||||||||||||||||||||||||||||||||||||:::
Db    427  HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleLysAlaProAlaGluAla  446

Qy    968  CCTGACTCCAAGTTCGTATGGTTATCATCGACGACGCCGCCGAGCCCCAATTCAAGGCT   1027
               ||:||||||||||||||||:||||||||||||||||||||||||||::::
Db    447  ProAspAlaLysValArgMetValIleThrGlyProGluAlaGlnPheLysAla  466

Qy   1028  CAGGGAAGAAATCATGGCAAACTCAAGGAGGAGAACTTCTTGGTCCAAGGAGGAAGTG   1087
               |||||||||:|||||||||||||||||||||||||||||:::
Db    467  GlnGlyArgIleThrGlyGlyLysLeuLysGluGluAsnPheValSerProLysGluVal  486

Qy   1088  AAGCTGGAGACCCACATACGTGTGCAGCATCAGCAGTTGCAGAGATTTGACGGACGTCAGTCATTGGCAAAGGT   1147
               ||||||||||||||||||:|||||||:||||||||||||||||||||||
Db    487  LysLeuGluLysHisIleArgValProSerPheAlaAlaGlyArgValIleIleGlyLysLysGlyGly  506

Qy   1148  GGAAAAACGGTGAACGAGTTGCAGAATTTGACGGACGTGAGGTGGTGAGGTGTAGTACCCAGAGAC   1207
               |||||:|||||||||||||||||:||||||:::|::||||||||:||:::
Db    507  GlyLysThrValAsnGluLeuGlnLeuGlnAsnLeuSerSerAlaGluValValValProProArgAsp  526

Qy   1208  CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGT   1267
               ||||||:::||||||:||||||||||:||:::|||||||||:::||
Db    527  GlnThrProAspGluAsnAspGlnValIleValValLysIleIleGlyHisPhePheAlaCys  546

Qy   1268  CAGATGGCTCAACGGAAGATCCGAAGATCCTGGCCCAGGTTAAGCAG---CAGCATCAG   1324
               |||:::||||||||||||||||||:::|||||||||||||||||||
Db    547  GlnIleAlaGlnArgLysIleGluGlnLeuIleLeuThrGlnValLysGlnHisGlnGln  566

Qy   1325  AAGGGA---CAGAGTAACCAGGCCCAGGCACGAGGCAAG   1360
               ||||     |||||||||||||||||||||||||||||
Db    567  LysAlaLeuGlnSerGlyProGlyGlnSerArgArgArgLys  579
```

RESULT 4
US-09-606-421B-348
; Sequence 348, Application US/09606421B
; Patent No. 6531315
GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579

Qy   908 CACATCAAACAGCTCTCCGGTTGCCAGCGCCTCCATCAGAATTGCACCACCCGAAACA 967
Db   133 ArgGlnAlaLeuAspLysLeuAsnGlyLysPheGlnLeuGluAsnPheThrLeuLysValAla 152

Qy   92 TACATCCCGATGAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCGAGGG------ 142
Db   153 TyrIleProAspGluMetAlaAlaGlnGlnArgProLeuGlnArgAsnProArgGlyArgArg 172

Qy   968 CCTGACTTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCT 1027
Db   427 HisIleIleLysGlnLeuSerArgPheAlaGlyMetAlaSerIleLysIleAlaProAlaGluAla 446

Qy   153 GGCTTTGGCTCTCGGGGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCC 202
Db   143 TyrIleProAspAlaLysValArgMetValIleIleThrGlyProGluAlaGlnPheLysAla 172

Qy 1028 CAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCAAGGAGGAGAGTG 1087
Db   173 GlyLeuGlyLysGlnArgGlySerSerArgGlnGlySerPro------GlySerValSer 189

Qy   467 AAGCAGCAGCAAGTGGACATCCCCTTGGCTCCTCGTGCCACCCAGTATGTGGGTGCC 262
Db   203 GlnGlyArgIleTyrArgGlyLysLysIleLeuLysGluGluLysProLysPheValSerProLysPro 486

Qy 1088 AAGCTGGAGACCCACATACGTGCCAGCATCAGCAGCTGGCCCGGGTCATTGGCAAAGGT 1147
Db   190 LysGlnLysProCysAspLeuProLeuValProThrGlnLeuLeuValProGlnPheValGlyAla 209

Qy   263 ATTATTGGCAAGGAGGGGCCCACCATCCGCAACATCAAAAACAGACCCAGTCCAAGATA 322
Db   263 LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyValIleGlyLysGly 506

Qy 1148 GGAAAAACGGTGACGAGTTGCGAGAATTGACGCGTGAGTGGTAGTACCAGAGAC 1207
Db   210 IleIleGlyLysGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229

Qy   507 GGTGCATAGGAACGACGCAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC 382
Db   323 AspValHisArgLysGluAlaAlaGlyAlaAlaGluLysGlnSerIleThrIleLeuSerThr 249

Qy 1208 CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAAATCATCCGGACATTTCTATGCCAGT 1267
Db   383 ProGluGlyThrSerAlaAlaCysLysSerProAlaCysLysLysGluAlaGln 269

Qy   547 GTAGTAGTAGGTGCTGCTCCTCCCCGCTTGAAGCAG---CAGCATCAG 1324
Db   443 AspIleLysPheThrGluGlyIleProLeuLysProLeuLeuAlaHisAsnAlaAsnPheValGly 502

Qy 1325 AAGGGA---CAGAGTAACCAGGCCCAGCCACCGAGGAAG 1360
Db   270 AspIleLysPheThrGluGlyIleProLeuLysProLeuLeuAlaHisAsnAlaAsnPheValGly 289

Qy   567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys 579
Db   503 CGTCTCATTGGCAAGGAAGGAGCGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAA 562
         ArgLeuIleGlyLysGluGlyLysTyrArgAsnLeuLysIleGluGlnIleAspThrAspThrLys 309

Db   563 ATCACCATCTCCTGGTTGCAAGACCTTCACCCTTACAACCCTGAGAGGACCATCACTGTG 622
         IleThrIleSerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrVal 329

Db   623 AAGGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGAAATAATGAAGAAAGTTCGGGAG 682
         LysGlyAlaIleGluAsnCysCysArgAlaGluGlnGluIleMetLysLysValArgGlu 349

Db   683 GCCTATGAGAATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGCCTGAAC 736
         SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuProLeuLeuAsn 369

Db   737 CTGGCTGCTGTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-------CCT 787
         LeuAlaAsnAlaLeuValLeuPheProThrSerGlyMetProProProThrSerGlyPro 389

Db   788 CCCAGCAGCGTTACTGGGGTGCTCCTATAGCTCCTTATGCAGGCTCCCGAGCAGGAG 847
         ProSerAlaMetThr------ProProTyrProGlnPheGluGlnSer---GluThrGlu 406

Db   848 ATGGTGCAGGTGTTTATCCCGCCCAGCCAGTGGGCGCCATCATCGCCAAGAAGGGCCAG 907
         ThrValHisGlnPheIleProAlaIle..:::|::||::||||||||||: 426

Db   908 CACATCAAACAGCTCTCCCGGTTGCCAGCGCCTCCATCAGAATTGCACCACCCGAAACA 967
         HisIleLysGlnLeuSerArgPheAlaGlyMetAlaSerIleLysIleAlaProAlaGluAla 446

Qy 1088 AAGCTGGAGACCCACATACGTGCCAGCATCAGCAGCTGGCCCGGGTCATTGGCAAAGGTCTCC 91
Qy   32 CGGGGAGCCATCATCGAAGTTGAATGGCCACCAGTTGGAGAACCAATGCCCTGAAGGTCTCC 91

Alignment Scores:
Pred. No.:   7.3e-146   Length: 579
Score: 1637.00   Matches: 336
Percent Similarity: 83.66%   Conservative: 43
Best Local Similarity: 74.17%   Mismatches: 58
Query Match: 52.64%   Indels: 16
DB: 4   Gaps: 9

US-09-270-437D-5 (1-1708) x US-09-643-597-176 (1-579)

```
Qy   487  LysLeuGluAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly  506

Qy  1148  GGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTGGTAGTACCAAGAGAC  1207
          ||||||||||||||||||||||| :::: :|||| ||||||::::|| || |||||||
Db   507  GGTLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAsp  526

Qy  1208  CAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGACATTTCTATGCCAGT  1267
          ||||||||||||||||||||||||| ||||||||||||||||||| || ||||
Db   527  GlnThrProAspGluAsnAspGlnValValIleValLysIleIleThrGlyHisPheTyrAlaCys  546

Qy  1268  CAGATGGCTCAACGGAAGATCCGAGACATCTCGAGACATCTCGGCCCAGGTAAGCAG---CAGCATCAG  1324
          ||||| |||||| |||||||||||||||| ::::|| || ||||||||
Db   547  GlnValAlaGlnArgLysIleGluIleLeuGluThrGlnIleLysGlnHisGlnGln  566

Qy  1325  AAGGGA---CAGAGTAACCAGGCCCAGGCACGGAGGAAG  1360
          ||| ||| |||:|||||||:|||
Db   567  LysAlaLeuGlnSerGlyProProGlnSerArgArgLys  579
```

RESULT 6
US-09-480-884A-176
; Sequence 176, Application US/09480884A
; Patent No. 6482597
GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-480-884A-176

Alignment Scores:
Pred. No.:   Length: 579
Score: 7.3e-146   Matches: 336
Percent Similarity: 1637.00   Conservative: 43
Best Local Similarity: 83.66%   Mismatches: 58
Query Match: 74.17%   Indels: 16
DB: 52.64%   Gaps: 9
                 4

US-09-270-437D-5 (1-1708) x US-09-480-884A-176 (1-579)

```
Qy    32  CGGGGAGCCATCATGAAGCTGAATGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC  91
          |||||    |||:::  ||||||||||||||||||||||| |||:::
Db    13  ArgGlnAlaLeuAspLysPheLeuAsnGlyPheGlnLeuGluAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla  152

Qy    92  TACATCCCCGATGACCAGATAGCA---CAGGGACCTGAGAATGGGCGCCGAGGG-----  142
          ||||||||||||:::||||| ||  ||||||||:::|||||||||
Db   153  TyrIleProAspGluMetAlaAlaGlnMetAlaAlaGlnLeuGlnAsnProLeuGlnLeuAsnProLeuGlnAlaArgGlyArgArg  172

Qy   143  GGCTTTGGCTCTCGGGTCAGCCCCCGGCCGGTCACCCTGGCACGGAGCGCCCCCAGCC  202
          ||||||||||||||||||||||||||| |||||||||:::||
Db   173  GlyLeuGlyAlaArgGlySerArgGlnGlyAlaArgGlySerPro------GlySerValSer  189

Qy   203  AAGCAGCAGCAAGTGGACATCGACACTCCCCTTCGGCTCCTGTGCGCCACCCAGTATGTGGGTGCC  262
          ||:::||   ||||:::  ||||||||||||:::  ||||||||||||
Db   190  LysGlnLysProCysAspLeuProLeuArgLeuLeuCysAlaThrGlnTyrValGlyAla  209

Qy   263  ATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACGACCCAGTCCAAGATA  322
          ||||||||||||||||||||||||||:::
Db   210  IleIleGlyLysGluGlyAlaThrIleArgAsnIleThrLysArgProSerProLysIle  229

Qy   323  CACGTCGCTACAGGACGAAGGAAGAACGCAGGTGCAGCTGCAAAAAGCCATCAGTGTGCACTCCACC  382
          ||||||||||:::||||  ||:::: |||
```

```
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-176

Alignment Scores:
Pred. No.:                             Length:         579
Score:                  1637.00        Matches:        336
Percent Similarity:     83.66%         Conservative:    43
Best Local Similarity:  74.17%         Mismatches:      58
Query Match:            52.64%         Indels:          16
DB:                     7.3e-146       Gaps:             9

US-09-270-437D-5 (1-1708) x US-09-542-615A-176 (1-579)

Qy   32 CGGGGAGCCATCATGAAGCTGATGAAGCTGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC   91
            |||  ||||::::  ||||||||||  ||| ||||||||||||||||||||||:::
Db  133 ArgGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnGlyPheLeuLysValAla       152

Qy   92 TACATCCCGATGAGCAGATAGCA---CAGGGACCTGTAGAATGGGCGCCGAGGG------      142
Db  153 TyrIleProAspGluMetAlaAlaGlnGlnAsnProLeuGlnProArgGlyArgArg       172

Qy  143 GGCTTTGCTCTCGGGGTCAGCCGCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCC        202
Db  173 GlyLeuGlyGlnArgGlySerSerArgGlnGlySerPro--------GlySerValSer     189

Qy  203 AAGCAGCAGCAAGTGGACATCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCC    262
Db  190 LysGlnLysProCysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAla   209

Qy  263 ATTATTGGCAAGGAGGGGGCCACCATCCGACAACATCACAAAACAGACCCAGTCCAAGATA  322
Db  210 IleIleGlyLysGluGlyGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIle 229

Qy  323 GACGTGCATAGGAGGAGGAAGGAAACGCAGGTGCAGTGAAAAGCCATCAGTGTGCACTCCACC 382
Db  230 AspValHisArgArgArgLysGluAsnAlaGlyValAlaLeuAlaGlyIleLysSerIleSerThr 249

Qy  383 CCTGAGGGCTGCTCCTCCGCTGTAAGCACCTTACCCTTACAACCCTGAGAGGGACCTAAG   442
Db  250 ProGluGlyThrSerArgAlaAlaCysLysLysSerIleLeuLeuGlnIleLeuMetHisLysGlnAsn 269

Qy  443 GACACCAAAACCGGCTGACGAGGTTCCCTGAAGATCCGAGATCTGGCCATAATAACTTTGTAGGG  502
Db  270 AspIleLysPhePheThrGluGlyIleLeuProLeuLysIleLeuLeuAlaHisAsnAsnPheValGly 289

Qy  503 CGTCTCATTGGCAAGGAAGGAGGCGAACCGAGAAGGTAGAGCAAGGAGACAAAA          562
Db  290 ArgLeuIleGlyLysGluGlyGlyArgAsnLeuLysLysLysLeuGluGluAsnThrAspThrLys   309

Qy  563 ATCACCATCTCCTCGTTGCAAGACCTTACCCTTACAACCCTGAGAGGACCATCACTGTG    622
Db  310 IleThrIleSerProLeuGlnAspLeuThrLeuTyrAsnProGluArgThrIleThrVal      329

Qy  623 AAGGGGGCCATCAGAGAATTGTTGCAGGGCCGAGCAGGAAATAATGAAGAAGTTCGGGAG  682
            ::||||          |||          ||::||||||||          ||||||
```

```
Db  330 LysGlyAsnValGluThrCysAlaLysAlaAlaGluGluGluIleMetLysLysIleArgGlu    349
Qy  683 GCCTATGAGAATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCTGGCCTGAAC        736
            :::||||||||::||:||||||        :::||||||:::||||||||||
Db  350 SerTyrGluAsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuProGlyLeuAsn        369
Qy  737 CTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCG-------CCT       787
            |||          |||||||||          |||||          :||||
Db  370 LeuAsnAlaLeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyPro      389
Qy  788 CCCAGCAGCGGTACTGGGGCTGCTCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAG       847
            |||          |||:::          |||          |||::: ||||||
Db  390 ProSerAlaMetThr------ProProTyrProGlnPheGlnGlnSer---GluThrGlu      406
Qy  848 ATGGTGCAGGTGTTATCCCGCCCAGGAGTGGCGCCATCATCGGCAAGAAGGGGCAG         907
            |||||::||||||          |||          ::|||          |||
Db  407 ThrValHisGlnPheIleProAlaLeuSerValGlyAlaIleIleIleIleGlyLysGlnGlyGln 426
Qy  908 CACATCAAACAGCTCTCCGTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACA        967
            |||||          |||          |||::||||||          :::
Db  427 HisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAla      446
Qy  968 CCTGACTCCCAAAGTTCGTATGTTATCATCATCACTGACGGCCAGAGCCCAATTCAAGGCT   1027
            |||:::          |||||||||          |||||||||          :::
Db  447 ProAspAlaLysValArgMetValIleIleThrAspGlyGlnSerProIleGlnGlyAla     466
Qy 1028 CAGGGAAGAATCTATGGCAAACTCAAGGACGGAGAACTTCTTGGTCCCAAGGAGGAAGTG    1087
            |||||          |||:::||          |||||||          |||||||
Db  467 GlnGlyArgIleTyrGlyLysLeuLysAspGlyGluLeuLeuValProLysGluGluVal     486
Qy 1088 AAGCTGGAGACCCCACATACGTGTGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGT   1147
            |||||          |||::::          |||          :::
Db  487 LysLeuGluIleAlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGly 506
Qy 1148 GGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTGGTAGTACCAAGAGAC     1207
            |||||          |||||          :::||          :::|||||
Db  507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgArgAsp     526
Qy 1208 CAGACCCCTGATGAGACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGT    1267
            |||          |||||          :::||||||          |||||||
Db  527 GlnThrProAspGluAsnAspGlnValValValLysIleIleThrGlyHisPheTyrAlaCys  546
Qy 1268 CAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCAGGTTAAGCAG---CAGCATCAG    1324
            |||||          |||::::          |||          :::|||
Db  547 GlnValAlaGlnArgLysIleArgAspIleLeuIleLeuThrGlnGlnAsnValLysGlnHisGlnGlnGln 566
Qy 1325 AAGGGA---CAGAGTAACCAGGCCCAGGACACGGAGGAAG                        1360
            ::::|||          |||||||
Db  567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys                         579
```

```
RESULT 8
US-09-606-421B-176
; Sequence 176, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
```

```
;
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-176
-
Alignment Scores:
Pred. No.:              7.3e-146      Length:           579
Score:                  1637.00       Matches:          336
Percent Similarity:     83.66%        Conservative:      43
Best Local Similarity:  74.17%        Mismatches:        58
Query Match:            52.64%        Indels:            16
DB:                     4             Gaps:               9

US-09-270-437D-5 (1-1708) x US-09-606-421B-176 (1-579)

Qy     32 CGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCC   91
             ||| |||||  ||||||| |||||||||::::|||||||||||||||||||||||
Db    133 ARGGlnAlaLeuAspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAla  152

Qy     92 TACATCCCGATGAGCAGATAGCA--CAGGGACCTGAGAATGGGCGCCGAGG------    142
             ||||||||||||||||       |||||
Db    153 TyrIleProAspGluMetAlaAlaGlnGlnAsnProLeuGlnLnProLeuGlnGlnArgGlyArgArg  172

Qy    143 GGCTTTGGCTCTCGGGGTCAGCCCCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCC  202
             ||| :::  ::
Db    173 GlyLeuGlyGlnArgGlySerArgGlnGlySerPro--------GlySerValSer  189

Qy    203 AAGCAGCAGCAGTGGACATCCCCTTCGGCTCCTGGTGCCACCCAGTATGTGGGTGCC  262
             ||||||||||  :::  :::|||||||||||||||||
Db    190 LysGlnLysProCysAspLeuProLeuLeuValProThrGlnPheValGlyAla  209

Qy    263 ATTATTGGCAAGGAGGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATA  322
             ||||||||||||||||||||
Db    210 IleIleIleGlyLysGluGluGlyProProSerAlaThrIleThrLysGlnThrGlnSerLysIle  229

Qy    323 GACGTGCATAGAAGGAAGAACCAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC  382
             ||||||||||||||||||||::::::||||:::::
Db    230 AspValHisArgArgLysGlnAsnAlaGlyValAlaGluLysSerIleSerValThr  249

Qy    383 CCTGAGGGCTGCTCCTCCGCTGTAAGATGATCTTGAAGATTATGCATAAAGAGGCTAAG  442
             ||| |||||::||::|||||||
Db    250 ProGluGlyThrSerAlaAlaCysLysSerIleLeuLysIleMetHisLysGluAlaGln  269

Qy    443 GACACCAAAACGGCTGACGAGGTTCCCCTGAAGATCCTGGCCCCATAATAACTTTGTAGGG  502
             |||||||||||||:::|||
Db    270 AspIleLysPheThrGluIleLeuProLeuLysIleLeuAlaHisAsnAsnPheValGly  289

Qy    503 CGTCTCATTGGCAAGGAAGGAGACGGAACCTGAGAAGGTAGAGCAAGAGATACCGAGACAAAA  562
             |||||||||||||||||||||:::::
Db    290 ArgLeuIleGlyLysGluGlyArgAsnLeuLysLysLysIleGluThrAspThrArgLys  309

Qy    563 ATCACCATCTCCTCGTTGCAAGACCTTACCCTTACAACCCTGAGAGGACCATCACTGTG  622
             |||||||||||||||||||
Db    310 IleThrIleSerProLeuGlnGluLeuThrLeuThrLysTyrAsnProGluArgThrIleThrVal  329

Qy    623 AAGGGGGCGTCTGAGAATTGTTGCAGGGCCGAGCAGGAAATAATGAAGAAGAGTTCGGGAG  682
             |||||||||   :::|
Db    330 LysGlyAlaSerGluAsnCysCysArgAlaGluGlnGluIleMetLysLysSerSerGlyGlu  349

Qy    683 GCCTATGAGAATGATGGGCTGCCATGAGC------TCTCACCTGATCCTGGCTGAAC  736
             :::||||||||:::||::
Db    350 SerTyrGluAsnAspMetGluLeuProCysAlaSerMetAsnLeuGlnAlaHisLeuProGlyLeuAsn  369

Qy    737 CTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCCGCAGTCCCGCCG-------CCT  787
             :::|||||       :::||||
Db    370 LeuAsnAlaLeuGlyLeuPheProThrGlySerGlyMetProProThrSerGlyTyrPro  389

Qy    788 CCCAGCAGCGGTTACTGGGGCTCCTTATGCAGGCTCCCTATAGCTCCTTATGCAGGGCTCCCGAGCAGGAG  847
             |||||||::::||        :::||||
Db    390 ProSerAlaMetThr-------ProProTyrProGlnPheGluGlnSer---GluThrGluGlu  406

Qy    848 ATGGTGCAGGTGTTTATCCCCGGCCACAGTGGGCGCCCATCATCGGCAAGAAGGGGCAG  907
             :::||||||       ::::|||
Db    407 ThrValHisGlnPheIleLeuProAlaLeuSerValAlaIleAlaIleLeuGlyLysLysGlyGlnGln  426
```

```
Qy    908 CACATCAAACAGCTCTCCGTTTGCCAGCGCCTCCATCAAGATTGACCACCCGAAACA   967
             |||||||||||||||||||||||||||
Db    427 HisIleLysGlnLeuSerArgPheAlaGlyAlaGlyArgPheLeuLysIleAlaProAlaGluAla   446

Qy    968 CCTGACTCCAAAGTTCGTATGGTATCATCACTGACCACCGCCAGAGGCCCAATTCAAGGCT   1027
             ||||:::::|||||||||||||||
Db    447 ProAspAlaLysValArgMetValIleIleThrGlyProGluGlnPheLysAla   466

Qy   1028 CAGGGAAGAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTG   1087
             |||||||||||||||:::||||
Db    467 GlnGlyArgArgIleTyrGlyLysLeuLysGluGluAsnPheVaSerProLysGluGluVal   486

Qy   1088 AAGTTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCCCGGGTCATTGGCAAAAGGT   1147
             ||| |||||||||:::::
Db    487 LysLeuGluLeuAlaHisIleArgValProSerPheAlaGlyArgValIleGlyLysLysGly   506

Qy   1148 GGAAAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGAC   1207
             |||||||||||||||||
Db    507 GlyLysThrValAsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgArgAsp   526

Qy   1208 CAGACCCCTGATGAGAACGACCAGTCATCGTGAAAATCATCGGACATTCTATGCCAGT   1267
             |||||||||||||||:::||||:::::|||
Db    527 GlnThrProAspGluAsnAspGlnSerSerCysGluLysSerSerGlyHisSerMetProSer   546

Qy   1268 CAGATGGCTCAACGAAGATCCGAGACATCCTGCCCAGGTTAAGCAG---CAGCATCAG   1324
             |||:::||||:::|||||||||||||||||
Db    547 GlnValAlaGlnArgLysIleGluGlnIleLeuLeuThrArgGlnValValLysGlnHisGlnGln   566

Qy   1325 AAGGGA--CAGAGTAACCAGGCCCAGGCACGGAGGAAG   1360
             |||:::      :::|||||||||||||
Db    567 LysAlaLeuGlnSerGlyProProGlnSerArgArgLys   579
```

RESULT 9
US-08-021-608D-10
; Sequence 10, Application US/08021608D
; Patent No. 5580760
GENERAL INFORMATION:
  APPLICANT: LEVENS, DAVID L., DUNCAN,
  APPLICANT: ROBERT C., AND AVIGAN, MARK I.
  TITLE OF INVENTION: NOVEL FUSE BINDING
  TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
  NUMBER OF SEQUENCES: 24
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: MORGAN & FINNEGAN
    STREET: 345 PARK AVENUE
    CITY: NEW YORK
    STATE: NEW YORK
    COUNTRY: USA
    ZIP: 10154
  COMPUTER READABLE FORM:
    MEDIUM TYPE: FLOPPY DISK
    COMPUTER: IBM PC COMPATIBLE
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: WORDPERFECT 5.1
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/021,608D
    FILING DATE: 22-FEB-1993
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:
    FILING DATE:
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: WILLIAM S. FEILER
    REGISTRATION NUMBER: 26,728
    REFERENCE/DOCKET NUMBER: 2026-4063
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (212) 758-4800
    TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 10:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 643

```
;   TYPE: Amino Acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Unknown
;   MOLECULE TYPE: Peptide/Protein
;   HYPOTHETICAL: No
;   ORIGINAL SOURCE:
;     ORGANISM: Human
;     CELL LINE: HL60
;   FEATURE:
;     OTHER INFORMATION:
;     OTHER INFORMATION: Amino Acid 148 (Xaa) is Met or Ile
US-08-021-608D-10

Alignment Scores:
Pred. No.:                            Length:         643
Score:                  1.15e-14      Matches:        140
Percent Similarity:     248.50        Conservative:    97
Best Local Similarity:  37.80%        Mismatches:     202
Query Match:            22.33%        Indels:         189
DB:                     7.99%         Gaps:            32
                        1

US-09-270-437D-5  (1-1708) x US-08-021-608D-10  (1-643)
```

Qy   98 CCCGAT---GAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGCTTTGGCTCT 154
             |||||| ::::::::|||||| |||||||:::||
Db   73 ProAspAlaLysLysValAla-----ProGlnAsn-------AspSerPheGlyThr 87

Qy   155 CGGGGTCAGCCCCGCCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGCAGCAA 214
             |||| :::||||||:::|||
Db   88 GlnLeuProProMetHis-------------------------GlnLnGlnArg 97

Qy   215 GTGGAGATCCCCCTTCGGCTCCTGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAG 274
             ||||||| ||||||||||| ||||||||||||||||
Db   98 SerValMetThrGluGluTyrLysValProAspGlyMetValGlyMetLysLeuArg 117

Qy   275 GAGGGGGCCACCATCCGCAACATCACAAAACAGACCCCAGTCCAAGATAGACGTGCATAGG 334
             ||| ::::::::: ||||||| ::: ||||| :::
Db   118 GlyGlyGluGlnIleSerArgIleGlnGluSerGlyCysLysIleGlnIle---Ala 136

Qy   335 AAGGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCTGAGGGCTGC 394
                                          :::|||||||||||||||||
Db   137 ProAspSerGlyTyrGlyLeuProGluArgSerCys***LeuThrGlyThrProGluSerVal 156

Qy   395 TCCTCCGCTTGTAAGATCATCTTGGAGATTATGCATAAA-------GAGGCTAAGGACACC 448
             |||||| :::|||||| :::|||||||| |||
Db   157 GlnSerAlaLysArgLeuLeuAspGlnIleValGluLysAspGlyArgProAlaProGlyPhe 176

Qy   449 AAAACGGCTGAC---------GAGGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTA 499
             ||| ::|||||::|||
Db   177 HisHisHisGlyAspGlyAsnAlaValGlnGluIleMetIleProAlaSerLysAla 196

Qy   500 GGGGGTCTCATTGGCAAGGAGAAGCGAACCTGAAGAAGGTAGACCGAAGATACCGAGACA 559
             |||||| :::||:::::::||||||||||||
Db   197 GlyLeuValIleGlyLysGlyLysGluProHisLeuLysGlnLeuLysGlnLeuArgGlyAlaGlyVal 216

Qy   560 AAAATCACCATCCTTCCTCGTTCCAAGACCTTACCCTTTACAACCCTGAGAGGACC---- 613
             |||:::   :::|||||||
Db   217 LysMetValMet-----IleGlnAsp-----------GlyProGlnAsnThrGlyAla 230

Qy   614 -------ATCACTGTGAAGGGGCCATCGAAATTGTTCAGGGGCCGAGCCAGGAAATA 664
             :::   :::   ::::::
Db   231 AspLysProLeuArgIleThrGlyAspTyrLysValGlnLnAlaLysGluMetVal 250

Qy   665 ATGAAGAAAGTTCCGGAG----------GCCTATGAGAATGATGTGGCTGCCATG 709
             |||||| |||
Db   251 LeuGluLeuIleArgAspGlnGlyGlyPheArgGluValArgAsnGluTyrGly----- 268

Qy   710 AGCTCTCACCTGATCCCCTGGCCTGCCTGCTGTGAGGTCTGTGTTTTTCCCAGCTTCATCC 769
             ||||| :::
Db   268 --- 268

Qy   770 AGCGCAGTGCCCGCGCGCCTCCCAGCAGCCGTTACTGGGCCGTCGCCCCATAGCTCCTTTATG 829
             ||| ::: |||
Db   268 --- 268

Db   269 ---------------------------------------SerArgIleGlyGly- 273

Qy   830 CAGGCTCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCAGGCAGTGGGCGCCATC 889
             |||||| |||||| |||||||||
Db   274 -----------AsnGluGlyIleLeuAspValProIleProArgPheAlaValGlyIleIleVal 289

Qy   890 ATCGGCAAGAAGGGCAGCACATCAAACAGCTCTCCCGGTTTGCCAGCGCCCTCATCAAG 949
             |||||||::: ||||||||:::|||
Db   290 IleGlyArgAsnGlyGluMetIleLysLysIleGlnAsnAspProAlaGlyValArgIleGln 309

Qy   950 ATTGCACCACCCGAA------ACACTGACTCCAAAGTTCTATGTGTTATCACTGGA 1003
             || ::: |||||||||||||||
Db   310 PheLysProAspAspGlyThrThrProGlu--------ArgIleAlaGlnIleIleThrGly 326

Qy   1004 CCGCCAGAG---GCCCAATTCAAGGCTCAG-------------------------- 1030
             ||||::: ||:::
Db   327 ProProAspArgCysGlnHisAlaAlaLaGluIleIleThrAspLeuLeuArgSerValGln 346

Qy   1031 -------------------GAAGAGATCTATGGCAAACTCAAGGAGGAGAG 1060
             :::
Db   347 AlaGlyAsnProGlyGlyProGlyGlyProGlyGlyArgGlyArgGlyGlnGlnAsn 366

Qy   1061 AACTTCTTTGGTCCCAAGGAGGAGAGTGAAGCTGGAGACCCACATACGTGCCAGCATCA 1120
             |||||||||::: ||||||||| |||
Db   367 TrpAsnMetGlyProProGlyGlyLeuGln-----GluPheAsnPheIleValProThrGly 385

Qy   1121 GCAGCTGCCGGGTCATTGGCAAAGGTGGAAAAACGTGAACGAGTTGCAGAATTTGACG 1180
             |||||| ||||||||||||:::|||
Db   386 LysThrGlyLeuIleIleGlyLysGlyGlyGluThrIleLysSerIleSerGlnGlnSer 405

Qy   1181 GCAGCTGAGGTGGTAGTACCAAGAGACAGGACCCCTGATGAGAACGACCAGGTC---ATC 1237
             |||||| |||||||||||||||:::|||||| :::
Db   406 GlyAlaAlaArgIleGluLeuGlnArgAsnProProProAsnAlaAspProAsnAlaAspProAsnAlaAspProAsnAlaAspMetLysLeu 425

Qy   1238 GTGAAAATCATCGGACAT-------------TTCTATGCCAGTCAGATGGCTCAACGGAAG 1285
             ||| |||||||
Db   426 PheThrIleArgGlyThrProGlnIleLeuAspTyrAlaArgGlnLeuLeuGluLeuGluLys 445

Qy   1286 AT------CCGAGACATCCT---GGCCCAGTTAAGCAGCAGCATCAGAGGACAGAG 1335
             ||| ||| :::|||
Db   446 IleGlyLeuGlyProValAsnProLeuGlyProProValProHisGlyVal--- 464

Qy   1336 TAACCAGGCCCAGGCACCGGA------------------GGAAGTGACCAGCCCCTC-- 1373
             |||||| |||
Db   465 ---ProGlyPro--HisGlyProProGlyProGlyProThrProMetGlyProThr 483

Qy   1374 -----------CCTGTCCCTTNGAGTCCAGA---------------------- 1394
             |||
Db   483 yrAsnProAlaProTyrArgAsnProGlyProProGlyProAlaProHisGlyProProGlyProAlaP 503

Qy   1395 -------------------------------CAACAACGGGCAGAGAAATCG 1413
                                            :::||||||||
Db   503 roTyrAlaProGlnIleTrpGlyAsnAlaTyrProHisTrpGlnAsnGlnGlnAlaProProA 523

Qy   1414 AGAGTGTGCTCTCCCCGGCCAGGCCTGAGAATGAGTGGAATCCGGACACNTGGCCGGG 1473
             ||| |||
Db   523 sp-----------ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAla 538

Qy   1474 CTGTAGATCAGTTGCCCACTTGCCACCTTGATTGAGAAAGATGTTCCAGTGAGGACCCTGATCTN 1533
             :::|||||| :::::::
Db   539 Tyr----------TyrAlaHisTyrTyrGlnGlnGlnAla 548

Qy   1534 TCAGCCCCAAACACCCACCCAATTGGC---CCAACACTGTNTGCCCCTCGGGGTGTCAGA 1590
             ||| |||||
Db   549 GlnProProProAlaAlaProAlaGlyValAlaProThrThrGlnThrArgGlnGlyGlyGly 568

Qy   1591 AATTNTAGGCGCAAGGCACTTTTAAACGTGAATTGTTTAAAGAAGCTCTCCAGGCCCCACC 1650
             :::   :: |||
Db   569 GlnLnAsnProAlaProAla 576

Qy   1651 AAGAGGGTGGATCACACC----------------TCAGTGGGAAGAAAAATA 1686
             :::|||:::   :::
Db   577 GlyGlnValAspTyrThrLysAlaTrpGluGluTyrTyrLysLysLysGlyGlnAlaVal 596

us-09-270-437d-5_1.rai

Qy   1687 AAATTCCTTCAGGT 1701
           ||||:|||:||
Db   597 ProAlaProThrGly 601

RESULT 10
US-08-726-160-10
; Sequence 10, Application US/08726160
; Patent No. 5734016
; GENERAL INFORMATION:
;   APPLICANT: LEVENS, DAVID L., DUNCAN,
;   APPLICANT: ROBERT C., AND AVIGAN, MARK I.
;   TITLE OF INVENTION: NOVEL FUSE BINDING
;   TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
;   NUMBER OF SEQUENCES: 24
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN & FINNEGAN
;   STREET: 345 PARK AVENUE
;   CITY: NEW YORK
;   STATE: NEW YORK
;   COUNTRY: USA
;   ZIP: 10154
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: FLOPPY DISK
;   COMPUTER: IBM PC COMPATIBLE
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WORDPERFECT 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/726,160
;   FILING DATE: 04-OCT-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/021,608
;   FILING DATE: 22-FEB-1993
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: WILLIAM S. FEILER
;   REGISTRATION NUMBER: 26,728
;   REFERENCE/DOCKET NUMBER: 2026-4063US1
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 758-4800
;   TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 10:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 643
;   TYPE: Amino Acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Unknown
;   MOLECULE TYPE: Peptide/Protein
;   HYPOTHETICAL: NO
;   ORIGINAL SOURCE:
;   ORGANISM: Human
;   CELL LINE: HL60
;   FEATURE:
;   OTHER INFORMATION:
;   OTHER INFORMATION: Amino Acid 148 (Xaa) is Met or Ile
US-08-726-160-10

Alignment Scores:
Pred. No.:   643
Score: 1.15e-14   Length: 643
Percent Similarity: 248.50   Matches: 140
Best Local Similarity: 37.80%   Conservative: 97
Query Match: 22.33%   Mismatches: 202
DB: 7.99%   Indels: 189
                1   Gaps: 32

US-09-270-437D-5 (1-1708) x US-08-726-160-10 (1-643)

Qy   98 CCCGAT--GAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGGCTTTCGCTCT 154
          ||||| ::::::||||||:|||||||| |||||||||||
Db   73 ProAspAlaLysLysValAla-----------ProGlnAsn----------AspSerPheGlyThr 87

Qy   155 CGGGGTCAGCCCCGCCAGGGCTCACCTGTGCGCAGGCGGGGGCCCCCAGCCAGCAGCAGCAA 214
          ||| :::

APPLICATION NUMBER: PCT/US94/01782
FILING DATE: 22-FEB-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/021,608
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAM S. FEILER
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4063PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 643
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide/protein
HYPOTHETICAL: No
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE: HL60
ORGANELLE:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
OTHER INFORMATION: Amino Acid 148 (Xaa) is Met or Ile

PCT-US94-01782-10

Alignment Scores:
Pred. No.:
Score: 1.15e-14   Length: 643
                                248.50   Matches: 140
Percent Similarity: 37.80%   Conservative: 97
Best Local Similarity: 22.33%   Mismatches: 202
Query Match: 7.99%   Indels: 189
DB: 5   Gaps: 32

US-09-270-437D-5 (1-1708) x PCT-US94-01782-10 (1-643)

```
Qy    98  CCCGAT--GAGCAGATAGCACAGGGACCTGAGAATGGCGCCGAGGGGGCTTTGGCTCT 154
          ||||||  ::::::::||||:||||        |||:|||
Db    73  ProAspAlaLysLysValAla------ProGlnAsn-------AspSerPheGlyThr 87

Qy   155  CGGGGTCAGCCCCGCCAGGGTCACCTGTGGCAGCGGGGGCCCAGCCAAGCAGCAGCAA 214
                                         :::|||||:|::::
Db    88  GlnLeuProProMetHis---------------------GlnGlnGlnArg 97

Qy   215  GTGGACATCCCCCTTCCCCTTCGGCTCCTGGTCCCCACCCAGTATGTGGTGCCATTATTGGCGAAG 274
                                      ||||||        |||||||||:
Db    98  SerValMetThrGluGluTyrLysValProAspGlyMetValGlyPheIleIleGlyIleGlyArg 117

Qy   275  GAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGG 334
                ||||        :::::::::: ::::::||||       :::
Db   118  GlyGlyGluGlnIleSerArgIleGlnGlnGluSerGlyCysLysLysGlnIle--Ala 136

Qy   335  AAGGAGAACGCAGGTGCAGTGAAAAAGCCATCAGTGTGACTCCCCTGAGGGCTGC 394
                     :::::||| :::                        :::
Db   137  ProAspSerGlyGlyLeuProGluArgSerCys***LeuThrGlyThrProGluSerVal 156

Qy   395  TCCTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAA-----GAGGGTAAGGACACC 448
                ||||||        :::|||| :::|||||        ||
Db   157  GlnSerAlaLysArgLeuLeuAspGlnIleValGluLysGlyAlaArgProAlaProGlyPhe 176
```

RESULT 11
PCT-US94-01782-10
; Sequence 10, Application PC/TUS9401782
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES
; APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
; APPLICANT: HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:

```
Db   386  LysThrGlyLeuIleIleGlyLysGlyGluThrIleLysSerIleSerGlnGlnSer 405
Qy  1181  GCAGCTGAGGTGGTGTAGTACCAGAGACACAGACCCCTGATGAGAACGACCAGGTC--ATC 1237
                |||    ::: :::||||::::       |||::: :::    :::
Db   406  GlyAlaArgIleGluValGlnArgAsnProProAsnAlaAspProAsnMetLysLeu 425
Qy  1238  GTGAAAATCATCGGACAT------------TTCTATGCCAGTCAGATGGCTCAACGGAAG 1285
                                        |||:||:
Db   426  PheThrIleArgGlyThrProGlnHisIleAspTyrAlaArgGlnLeuIleGluGluLys 445
Qy  1286  AT-----CCGAGACATCCT--GGCCCAGTTAAGCAGCAGCATCAGAAGGGACAGAG 1335
                   |||     :::|||     |||
Db   446  IleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHisGlyVal--- 464
Qy  1336  TAACCAGGCCCAGGCACCGA-------------GGAAGTGACCAGCCCCTC------- 1373
                ||||||||| -||||||       |||:::
Db   465  ---ProGlyPro--HisGlyProProGlyProGlyThrProMetGlyProT 483
Qy  1374  ------CCTGTCCCTTNGAGTCCAGGA-------------------- 1394
          ||:
Db   483  yrAsnProAlaProTyrAsnProGlyProProGlyProAlaProHisGlyProProAlaP 503
Qy  1395  --------------------------CAACACGGGCAGAAATCG 1413
                                     |||||||:::|||
Db   503  roTyrAlaProGlnGlyTrpGlyTrpGlyTyrAsnAlaTyrProHisTrpGlnGlnGlnAlaProProA 523
Qy  1414  AGAGTGTGCTCTCCCCGCAGGCCTGAGAATGAGTGGGAATCCGGGACACNTGGGCCGGG 1473
                      ||||||        |||
Db   523  sp--------ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAla 538
Qy  1474  CTGTAGATCAGGTTTGGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGATCTN 1533
                ::|||||        :::|||
Db   539  Tyr-------TyrAlaHisTyrTyrGlnGlnGlnAla------------- 548
Qy  1534  TCAGCCCCAAACACCCACCCAATTGGC--CCAACACTGTNTGCCCCTCGGGGTGTCAGA 1590
                ||||||       |||  |||||
Db   549  GlnProProAlaAlaProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGly 568
Qy  1591  AATTNTAGGCGCAAGGCACTTTTAAACGTGGATTGTTTAAAGAAGCTCTCCAGGGCCCACC 1650
Qy  1651  AAGAGGGTGGATCACACC------------------------TCAGTGGGAAGAAAAATA 1686
                ::|||||||:::||        |||||:|:::|||
Db   569  Asp----------------------------GlnGlnAsnProAlaProAla 576
Db  1651                                              ||:
Qy  1687  AAATTTCCTTCAGT 1701
          |||:|||
Qy  577  GlyGlnValAspTyrThrLysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaVal 596
Db   597  ProAlaProThrGly 601
```

Qy   449 AAAACGGCTGAC--------GAGGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTA 499
                |||| ||| :::
Db   177 HisHisGlyAspGlyProGlyAsnAlaValGlnGluIleMetIleProAlaSerLysAla 196

Qy   500 GGGCGTTCATTGGCAAGGAACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACA 559
                :::|||||||||| ||||||:::||
Db   197 GlyLeuValIleGlyLysGlyGlyThrIleLysGlyGluThrIleLysGlnLeuArgAlaGlyAlaGlyLysVal 216

Qy   560 AAAATCACCATCTCCTGTTGCAAGACCTTACCCTTTACAACCTGAGAGACC------- 613
                ||||::: ::: |||||||
Db   217 LysMetValMet-----GlyProGlnAsnThrGlyAla 230

Qy   614 ------ATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGAAATA 664
                :::   ::: ||| |||
Db   231 AspLysProLeuArgIleThrGlyAspProTyrLysLysValGlnGlnAlaLeuLysGluMetVal 250

Qy   665 ATGAAGAAGTTCGGGAG-------GCCTATGAGAATGATGTGGCTGCCCATG 709
                ::::: |||| ||:::|||||
Db   251 LeuGluLeuIleArgAspGlnGlyGlyPheArgGluValArgAsnGluTyrGly------ 268

Qy   710 AGTCTCACCTGATCCCTGGCCTGAACCTGGCTGCGTAGGTCTTTCCCAGTCTTCATCC 769
                |||||||||||
Db   268   268

Qy   770 AGCGCAGTCCCCGCCGCCTCCCAGCAGCGTACTGGGGCTGCTCCCTATAGCTCCTTTATG 829
                :::::::||::: |||
Db   269   SerArgIleGlyGly- 273

Qy   830 CAGGCTCCCGAGCAGGAGGATGTGCAGGTGTGTTATCCCCGAGCAGTGGGCGCCATC 889
                ||||||| ||::: |||||||
Db   274 ------AsnGluGlyIleLeuAspValProIleLeuProArgPheAlaValGlyIleValVal 289

Qy   890 ATCGGCAAGAAGGGGCCAGCACATCAAACAGCTCTCCCGGTTTGCCAGCGCCTCCATCAAG 949
                ||||::: ||::: |||:::
Db   290 IleGlyArgAsnGlyGluMetIleLeuLysLysIleIleLysLysLysIleIleAsnAspAlaGlyValAlaGlyIleValAlaArgIleGln 309

Qy   950 ATTGCACACCCGAA-----ACACTGACTCCAAAGTTCGTATGTGTTATCATCACTGGA 1003
                ||||||||||||
Db   310 PheLysProAspAspProGlyThrThrProGlu-------ArgIleAlaGlnIleThrGly 326

Qy   1004 CCGCCAGAG--GCCCAATTCAAGGCTCAG----- 1030
                ||||||::: |||
Db   327 ProProAspArgCysGlnHisAlaLeuAlaGluIleIleThrAspLeuLeuArgSerValGln 346

Qy   1031 ------------GGAAGAATCATGCGAACTTCAAGGAGGAG 1060
                ||||||||| ::: :::
Db   347 AlaGlyAsnProGlyProGlyProGlyTyrArgGlyArgGlyArgGlyGlnGlyAsn 366

Qy   1061 AACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCA 1120
                |||:::
Db   367 TrpAsnMetGlyProProGlyGlyLeuGln---GluPheAsnProPheIleValProThrGly 385

Qy   1121 GCAGCTGGCGCGGGTCATTGGCAAAAGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACG 1180
                ||||:::|||||||||||:::
Db   386 LysThrGlyAsnProGlyLeuIleLeuGlyLysGlyGlyGlnThrIleLysSerIleSerGlnGlnSer 405

Qy   1181 GCAGCTGAGGTGGTAGTACCAAGAGAACGAGACCAGGTC---ATC 1237
                |||| ::: :::
Db   406 GlyAlaAlaArgIleGluLeuGlnArgAsnProProAsnAlaAspProAsnMetLysLeu 425

Qy   1238 GTGAAAATCATCGGACAT-----------TTCTATGCCAGTCAGATGGCTCAACGGAAG 1285
                ||||:::
Db   426 PheThrIleArgGlyThrProGlnIleAsnIleLeuAspTyrAlaArgGlnIleLeuIleGluGluLys 445

Qy   1286 AT-------CCGAGACATCCT---GGCCCAGGTTAAGCAGCAGCATCAGAAGGGACAGAG 1335
                || :::|||
Db   446 IleGlyGlyProValAsnProLeuGlyProValProLeuProGlyProHisGlyProVal--- 464

Qy   1336 TAACCAGGCCCAGGCCACGGA-----GGAAGTGACCAGCCAGCCCCTC------- 1373
                ||||||||| |||||||
Db   465 ---ProGlyPro-HisGlyProGlyProGlyProGlyProMetGlyProT 483

Qy   1374 -----CCTGTCCCTTNGAGTCCAGGA 1394

Qy   483 yrAsnProAlaProTyrAsnProGlyProGlyProAlaProHisGlyProProGlyProProAlaP 503
                ||| |||||||
Db   1395 --------------------------CAACAACGGGCCAGAAATCG 1413

Qy   503 roTyrAlaProGlnGlyThrGlyAsnAlaTyrProHisTrpGlnGlnAlaProProA 523
                |||||||||||:::|||
Db   1414 AGAGTGTGCTCTCCCGGCAGGCTGAGAATGGGAATCCGGACANTGGGCCGGG 1473

Qy   523 sp-----------ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAla 538
                :: |||||||
Db   1474 CTGTAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAGGAACCCTGATCTN 1533

Qy   539 Tyr-----------------TyrAlaHisTyrTyrGlnGlnGlnAla--- 548
                :::   ::::::
Db   1534 TCAGCCCCAAACACCACCCCAATTGGC--CCAACACTGTNTGCCCCTCGGGGTGTCAGA 1590

Qy   549 GlnProProProAlaProAlaGlyAlaProThrThrGlnThrProGlnGly 568
                |||||||:::|||
Db   1591 AATTNTAGCGCAAGGCACTTTAAACGTGGATTGTTTAAAGAAGCTCTCCAGGCCCCACC 1650

Qy   569 Asp-----------------GlnGlnAsnProAlaProAla 576
                :::
Db   1651 AAGAGGGTGGATCACACC------------TCAGTGGGAAGAAAAAATA 1686

Qy   577 GlyGlnValValAspProTyrThrLysValAlaTrpGluGluTyrTyrLysLysGlyMetGlyAlaValAlaVal 596
                :::||||||
Db   1687 AAATTTCCTCAGGT 1701

Qy   597 ProAlaProThrGly 601
                ||:::
Db   1687   1701

RESULT 12

Alignment Scores:
Pred. No.:   5.86e-15   Length: 49
Score: 245.00   Matches: 49
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 7.88%   Indels: 0
DB: 3   Gaps: 0

US-09-270-437D-5 (1-1708) x US-09-261-855-18 (1-49)

Qy   476 ATCCTGGCCCATAATAACTTTGTAGGGCGTTCATTGGCAAGGACGGAACCTGAAG 535
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1   IleLeuLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyLysLeuLys 20

Qy   536 AAGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTGTTGCAAGACCTTACCCTT 595
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21   LysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeuThrLeu 40

Qy   596 TACAACCCTGAGAGGACCATCACTGTG 622
                ||||||||||||||||||||||
Db   41   TyrAsnProGluArgThrIleThrVal 49

RESULT 13

US-08-021-608D-2
; Sequence 2, Application US/08021608D
; Patent No. 5580760
; GENERAL INFORMATION:
;   APPLICANT: LEVENS, DAVID L., DUNCAN,
;   APPLICANT: ROBERT C., AND AVIGAN, MARK I.
;   TITLE OF INVENTION: NOVEL FUSE BINDING
;   TITLE OF INVENTION: PROTEIN AND cDNA THEREFOR
;   NUMBER OF SEQUENCES: 24
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN & FINNEGAN
;   STREET: 345 PARK AVENUE
;   CITY: NEW YORK
;   STATE: NEW YORK
;   COUNTRY: USA
;   ZIP: 10154
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: FLOPPY DISK
;   COMPUTER: IBM PC COMPATIBLE
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WORDPERFECT 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/021,608D
;   FILING DATE: 22-FEB-1993
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: WILLIAM S. FEILER
;   REGISTRATION NUMBER: 26,728
;   REFERENCE/DOCKET NUMBER: 2026-4063
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (212) 758-4800
;   TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 644
;   TYPE: Amino Acid
;   STRANDEDNESS: Single
;   TOPOLOGY: Unknown
;   MOLECULE TYPE: Peptide/Protein
;   HYPOTHETICAL: No
;   ORIGINAL SOURCE:
;   ORGANISM: Human
;   CELL LINE: HL60
;   FEATURE:
;   OTHER INFORMATION:
;   OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile

US-08-021-608D-2

Alignment Scores:
Pred. No.:
Score: 2.46e-14
Score: 245.00
Percent Similarity: 36.46%
Best Local Similarity: 22.39%
Query Match: 7.88%
DB: 1

Length: 644
Matches: 148
Conservative: 93
Mismatches: 219
Indels: 202
Gaps: 33

US-09-270-437d-5 (1-1708) x US-08-021-608D-2 (1-644)

```
;       CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/021,608
;       FILING DATE: 22-FEB-1993
;       CLASSIFICATION: 435
;       ATTORNEY/AGENT INFORMATION:
;       NAME: WILLIAM S. FEILER
;       REGISTRATION NUMBER: 26,728
;       REFERENCE/DOCKET NUMBER: 2026-4063US1
;       TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (212) 758-4800
;       TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;       LENGTH: 644
;       TYPE: Amino Acid
;       STRANDEDNESS: Single
;       TOPOLOGY: Unknown
;       MOLECULE TYPE: Peptide/Protein
;       HYPOTHETICAL: No
;       ORIGINAL SOURCE:
;       ORGANISM: Human
;       CELL LINE: HL60
;       FEATURE:
;       OTHER INFORMATION:
;       OTHER INFORMATION:    Amino Acid 149 (Xaa) is Met or Ile
US-08-726-160-2

Alignment Scores:
Pred. No.:                              2.46e-14
Score:                                  245.00
Percent Similarity:         36.46%          Length:       644
Best Local Similarity:      22.39%          Matches:      148
Query Match:                7.88%           Conservative: 93
DB:                         1               Mismatches:   219
                                            Indels:       202
                                            Gaps:         33
```

US-09-270-437D-5 (1-1708) x US-08-726-160-2 (1-644)

```
Qy    71 AACCATGCCCTGAAGTCTCCTACATCCCCGATGAGCAGATAGCACAG-----GGACCT 124
         ||| |||||| |||| ||   :::                          ||:::
Db    28 AsnAspAlaPheLysAspAlaLeuGlnArgAlaArgGlnIleAlaAlaLysIleGlyGly  47

Qy   125 GAGAATGGG---------CGCCGAGGGGCTTGGCTCCGGGGTCAGCCCCCCAGGGC 175
         ||:                                    ||||||||||||
Db    48 AspAlaGlyThrSerLeuAsnSerAsnAspTyrGlyTyrGlyTyrGlyGlnIlelysArg-------  65

Qy   176 TCACCTGTGGCAGCGGGG---------                                193
         |||:::                                                    ||||||
Db    66 ---ProLeuGluAspGlyAspGlnProAspAlaLysLysValAlaProGlnAsnAspSer  84

Qy   194 --------------GCCCCAGCCAAGCAGCAGCAAGTGGACATCCCCTT-----CGG 232
         ||                                  :::
Db    85 PheGlyThrGlnLeuProProMetHisGlnLeuGlnGlnAsnSerArgSerValMetThrGluGlu 104

Qy   233 CTCTTGGTGCCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGC 292
         :::                                  |||:|||: ||
Db   105 TyrLysValProAspGlyMetValGlyPheIleIleGlyArgGlyGlyGlnIleSer 124

Qy   293 AACATCACAAAACAGACCCCAGTCCAAGATAGACGTGCATAGGAGGAGACGACAGGTGCA 352
         ||                                  :::
Db   125 ArgIleGlnGlnGluSerGlyCysLysIleGlnIle---AlaProAspSerGlyGlyLeu 143

Qy   353 GCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTTGTAAGATG 412
         ::::                                 :::
Db   144 ProGluArgSerCys***LeuThrGlyThrProGluSerValGlnSerArgLysArgLeu 163

Qy   413 ATCTTGGAGATTATGCATAAA----------GAGGCTAAGGACCAACCAAAACGGCTGAC------ 460
         :::: ||:::: | ||                      |||::
Db   164 LeuAspGlnIleValGlnLysGlyValGlyArgProAlaProGlyPheHisGlyAspGlyPro 183

Qy   461 ---GAGGTTCCCCTGAAGATCCTGCCCCATAATAACTTGTGAGGGCGTCTCATTGGCAAG 517
         :::|:|||| :::                         ||:|||||||
Db   184 GlyAsnAlaValGlnIleMetIleProAlaSerLysAlaGlyAlaLeuValIleGlyLys 203
```

```
Qy  1139 GGCAAAGGTGGAAAAACGGTGACGAGTTGCAGGCAGCTGAGGTGGTAGTA 1198
         ||||||||||| ||:::                    :::
Db   393 GlyLysGlyGlyGluThrIleLysSerIleSerGlnInSerGlyAlaArgIleGluLeu 412

Qy  1199 CCAAGAGACCAGACCCCTGATGAGAACGACCAGGTC--ATCGTGAAAATCATCGGACAT 1255
         ||:::  |||||||                       |||  :::
Db   413 GlnArgAsnProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr 432

Qy  1256 ----------TTCTATGCCAGTCAGATGGCTCAACGGAAGAT------CCGAGACAT 1296
         :::                ||||||   :::             ||||
Db   433 ProGlnGlnIleAspTyrAlaArgGlnLeuIleGluGluLysIleGluGlyProValAsn 452

Qy  1297 CCT--GGCCCAGTTAAGCAGCAGCATCAGAAGGGACAGAGCAGACCCCAGGCACG 1353
         :::  |||||                ||||||||    ||||:::
Db   453 ProLeuGlyProProValProHisGlyValPro-----ProGlyPro--HisG 470

Qy  1354 GA--------GGAAGTGACCAGGCCCTC------CCTGTCCCTTNGA 1386
         |||||             ||||||||||              ||||
Db   470 lyProProGlyProProGlyThrProMetGlyProTyrAsnProAlaProTyrA 490

Qy  1387 GTCCAGGA-------------------------------------------------- 1394
         ::||||||
Db   490 snProGlyProProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyT 510

Qy  1395 ------------------------CAACAACGGGCAGAAATCCAGAGTGTGCTCTCCCGG 1431
         :::                     ||||||:::|:
Db   510 rpGlyAsnAlaTyrProHisTyrProGlnGlnAlaProProAsp-----------ProA 526

Qy  1432 CAGGCCTGAGAATGAGATGGAATCCGGGACACNTGGCGCCGGGCTGTAGATCCAGGTTTGCC 1491
         ||||||                               :::
Db   526 laLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAlaTyr------TyrAla 542

Qy  1492 CACTTGATTGAGAAGATGTTCCAGTGAGATCNTCAGCCCCAAACACCAC 1551
         ||                          :::::::
Db   543 HisTyrTyrGlnGlnGlnAla--------------GlnProProProAlaAla 555

Qy  1552 CCAATTGGC--CCAACACTGNTGCCCCTGCGGGGTGTCAGAAATNTAGCGCAAGGCAC 1608
         ||||||     |||||
Db   556 ProAlaGlyAlaProThrThrThrGlnThrThrAsnGlyIleGlnGlyLysAsp---- 570

Qy  1609 TTTAAACGTGGATTGTTTAAGAAAGTCTCCAGGCCCCACCCAAGAGGGTGGATCACACC 1668
         :::    ||:|||||:::||||::
Db   571 -----------GlnGlnAsnProAlaProAlaGlyIleGlnnValAspTyrThr 583

Qy  1669 ------------------------TCAGTGGGAGAGAAAAATAAAATTTCCTTCAGGT 1701
         :::|||||:::  :::
Db   584 LysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaValProAlaProThrGly 602
```

RESULT 14
US-08-726-160-2
; Sequence 2, Application US/08726160
; Patent No. 5734016
; GENERAL INFORMATION:
;   APPLICANT: LEVENS, DAVID L., DUNCAN,
;   APPLICANT: ROBERT C., AND AVIGAN, MARK I.
;   TITLE OF INVENTION: NOVEL FUSE BINDING
;   TITLE OF INVENTION: PROTEIN AND cDNA THEREFOR
;   NUMBER OF SEQUENCES: 24
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORGAN & FINNEGAN
;   STREET: 345 PARK AVENUE
;   CITY: NEW YORK
;   STATE: NEW YORK
;   COUNTRY: USA
;   ZIP: 10154
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: FLOPPY DISK
;   COMPUTER: IBM PC COMPATIBLE
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: WORDPERFECT 5.1
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/726,160
;   FILING DATE: 04-OCT-1996

Qy 1395 ---------------------------------CAACAACGGGCAGAAATCGAGAGTGTGCTCTCCCCGG 1431
                                           |||||:::   :::
Db   510 rpGlyAsnAlaTyrProHisTrpGlnGlnGlnAlaProProAsp-----------ProA 526

Qy 1432 CAGGCCTGAGAATGAGTGGGAATCCGGGACANTGGGCCGCCGGCTGTAGATCAGGTTTGCC 1491
                                           |||||
Db   526 laLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAlaTyr---------TyrAla 542

Qy 1492 CACTTGATTGAGAAAGATGTTCCAGTGAGAAACCCTGATCNTCAGCCCCAAACACCCAC 1551
                                           :::::::
Db   543 HisTyrTyrGlnGlnGlnAla--------------------GlnProProAlaAla 555

Qy 1552 CCAATTGGC--CCAACACTGTNTGCCCTCGGGGTGTCAGAAAATNTAGGCGCAAGGCAC 1608
                                           |||||
Db   556 ProAlaGlyAlaProThrThrGlnThrThrAsnGlyGlnGlyAsp-   570

Qy 1609 TTTTAAACGTGGATTGTTGTTAAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACC 1668
                                           :::   :::|||
Db   571 ----------------------GlnGlnAsnProAlaProAlaGlyGlnValAspTyrThr 583

Qy 1669 --------------------------------TCAGTGGGAAGAAAAATAAAATTCCTTCAGGT 1701
                                           :::   :::
Db   584 LysAlaTrpGluGluTyrTyrLysMetGlyLysMetGlyGlnAlaValProAlaProThrGly 602

RESULT 15
PCT-US94-01782-2
; Sequence 2, Application PC/TUS9401782
GENERAL INFORMATION:
   APPLICANT: THE GOVERNMENT OF THE UNITED STATES
   APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
   APPLICANT: HEALTH AND HUMAN SERVICES
   TITLE OF INVENTION: NOVEL FUSE BINDING
   TITLE OF INVENTION: PROTEIN AND cDNA THEREFOR
   NUMBER OF SEQUENCES: 10
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: MORGAN & FINNEGAN
      STREET: 345 PARK AVENUE
      CITY: NEW YORK
      STATE: NEW YORK
      COUNTRY: USA
      ZIP: 10154
   COMPUTER READABLE FORM:
      MEDIUM TYPE: FLOPPY DISK
      COMPUTER: IBM PC COMPATIBLE
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: WORDPERFECT 5.1
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: PCT/US94/01782
      FILING DATE: 22-FEB-1994
      CLASSIFICATION:
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: U.S. 08/021,608
      FILING DATE: 22-FEB-1993
   ATTORNEY/AGENT INFORMATION:
      NAME: WILLIAM S. FEILER
      REGISTRATION NUMBER: 26,728
      REFERENCE/DOCKET NUMBER: 2026-4063PCT
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: (212) 758-4800
      TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 644
      TYPE: Amino Acid
      STRANDEDNESS: Single
      TOPOLOGY: Unknown
      MOLECULE TYPE: Peptide/Protein
      HYPOTHETICAL: NO
   ORIGINAL SOURCE:
      ORGANISM: Human
      STRAIN:
      INDIVIDUAL ISOLATE:
      DEVELOPMENTAL STAGE:

Qy   518 GAAGGACCGAACCTGAGAAGGTAGACGAAGATACCGAGACAAAATCACCATCTCTCG 577
                                           :::::::::::::   :::
Db   204 GlyGlyGluThrThrLeuLysGlnLeuGlnLeuGluArgAlaGlyValLysMetValMet------ 221

Qy   578 TTGCAAGACCTTACCCTTTACAACCCTGAGAGGACC------------ATCACTGTG 622
                                           |||:||||| :::   :::
Db   222 IleGlnAsp-----------GlyProGlnAsnThrGlyAlaAlaAspLysProLeuArgIle 237

Qy   623 AAGGGGGCCATCCAGAATTGTTCCAGGGCCAGGAAATAATGAGAGAAGTTCGGGAG 682
                                           ||||||::::::   :::::::
Db   238 ThrGlyAspProTyrLysValGlnGlnAlaLysGluMetValLeuGluLeuIleArgAsp 257

Qy   683 ------------GCCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGATCCCT 727
                                           |||::::
Db   258 GlnGlyGlyPheArgGluValArgAsnGluTyrGly----------------------- 269

Qy   728 GGCCTGAACCTGCTGCTGTGTAGGTCTTTCCCAGCTTCATCCAGCCGCAGTCCCGCCGCCT 787
                                           :::   :::|||
Db   269 -- 269

Qy   788 CCCAGCAGCCGGTTACTGGGGCGTCCTCCCTATAGCTCCTTATGCAGGCTCCCCAGCAGGAG 847
                                           |||::||
Db   270 ---SerArgIleGlyGly------------------------------------AsnGlu 276

Qy   848 ATGTGTGCAGGTGTTTATCCCGCCCAGGCGCGCCATCATCGGCCAAGAAGGGGCAG 907
                                           |||||| :::||::   :::|||
Db   277 GlyIleAspValProIleProArgPheAlaValAlaGlyIleValIleGlyArgAsnGlyGlu 296

Qy   908 CACATCAAACAGTCTCCCGGTTGCCAGCGGCCTCCATCAAGATTGCACCACCCGGA--- 964
                                           |||||| :::|||
Db   297 MetIleLysLysLysGlnAsnAspAlaGlyValArgIleGlnIlePheLysProAspPspGly 316

Qy   965 --ACACCTGACTCCAAAGTTCGTATGTTATCATCATCTGGACCGCCAGAG--GCCCAA 1018
                                           |||||| :::   :::|||
Db   317 ThrThrProGlu--------ArgIleAlaGlnIleThrGlyProProAspArgCysGln 333

Qy 1019 TTCAAGGCTCAG-- 1030
                                           |||::|
Db   334 HisAlaAlaGluIleLeuThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGly 353

Qy 1031 ------------GGAAGAATCATGGCAAACTCAAGGAGGAGGAGAACTTCTTTGGTCCCAAG 1078
                                           |||:||
Db   354 ProGlyProGlyGlyArgGlyArgGlyGlyArgGlyValGlnGlyAsnTrpAsnMetGlyProPro 373

Qy 1079 GAGGAAGTGAAGCTGGAGACCCCACATACGTGTGCCAGCATCACGCAGCTGGCCCGGGTCATT 1138
                                           |||||| :::|||
Db   374 GlyGlyLeuGln---GluPheAsnPheIleValProThrThrGlyLysThrGlyLeuIleIle 392

Qy 1139 GGCAAAGGTGGAAAAACCGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTA 1198
                                           |||||||||
Db   393 GlyLysGlyGlyGluThrThrLeuLysSerIleSerGlnLeuSerGlyAlaArgIleGluLeu 412

Qy 1199 CCAAGAGACCAGACCCCCGATGAGAACGACCAGGTC--ATCGTGAAAATCATCGGACAT 1255
                                           |||:::   :::|||
Db   413 GlnArgAsnProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr 432

Qy 1256 ----------TTCTATGCCAGTCAGATGGCTCAACCGAAGAT------CCGAGACAT 1296
                                           |||||| :::
Db   433 GlyLysGlyGlyGluThrThrLeuLysGlnLeuGlnLeuGluGlyLysGlyProValAsn 452

Qy 1297 CCT---GGCCCAGGTTAAGCAGGTATCCAGAAGGGACAGAGTAACCAGGCCCAGGCACG 1353
                                           ||||| :::||||
Db   453 ProLeuGlyProValProProValProHisGlyProHisGlyVal------ProGlyPro--HisG 470

Qy 1354 GA------------GGAAGTGACCAGCCCCTC------------CCTGTCCCTTNGA 1386
                                           |||::
Db   470 lyProProGlyProProGlyThrProMetGlyProTyrAsnProAlaProTyrA 490

Qy 1387 GTCCAGGA-- 1394
                                           |||:||
Db   490 snProGlyProProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyT 510

```
;    HAPLOTYPE:
;    TISSUE TYPE:
;    CELL TYPE:
;    CELL LINE:             HL60
;    ORGANELLE:
;    FEATURE:
;    NAME/KEY:
;    LOCATION:
;    IDENTIFICATION METHOD:
;    OTHER INFORMATION:
;    OTHER INFORMATION:     Amino Acid 149 (Xaa) is Met or Ile
PCT-US94-01782-2

Alignment Scores:
Pred. No.:
Score:                    2.46e-14
                          245.00
Percent Similarity:       36.46%
Best Local Similarity:    22.39%
Query Match:              7.88%
DB:                       5

Length:       644
Matches:      148
Conservative: 93
Mismatches:   219
Indels:       202
Gaps:         33

US-09-270-437D-5 (1-1708) x PCT-US94-01782-2 (1-644)
```

```
Qy    71   AACCATGCGCCTGAAGGTCTCCTACATCCCGATGAGCAGATGCACAG------GGACCT    124
           |||  ||  |||    :::                                :::
Db    28   AsnAspAlaPheLysGlyValAspAlaLeuGlnIleLeuArgGlnIleAlaAlaLysIleGlyGly   47

Qy    125  GAGAATGGG---------CGCCCAGGGGCTTTGGCTCTCCGGGGTCAGCCCGCCAGGGC    175
           :::   |||         |||:::|||:||     :::|||
Db    48   AspAlaAlaGlyThrSerLeuAsnSerAsnAspTyrGlyTyrGlyValGlyGlnLysArg------   65

Qy    176  TCACCTGTGGCAGCGGGG---------                                    193
           |||  ||| ||||
Db    66   ---ProLeuGluAspGlyAspGlnProAspAlaLysLysValAlaProGlnAsnAspSer    84

Qy    194  ------------GCCCCAGCCAAGCAGCAGCAGCAGGACCATCCCCTT------CGG      232
                       :::|||:|||:::|||
Db    85   PheGlyThrGlnLeuProProMetHisGlnHisGlnSerArgSerValMetThrGluGlu   104

Qy    233  CTCCTGGTGCCCACCCCAGTATGTGGGTGCCATTAGTGGCAAGGAGGAGGGGCCACCATCCGC   292
           :::|||   |||||||:::|||          |||:::|||:::||||||||
Db    105  TyrLysValProAspGlyMetGlyMetValGlyPheIleLeuGluLeuTyrGlyGlyGlyGluGlnIleIleSer   124

Qy    293  AACATCACAAAACAGACCCCAGTCCAAGATAGACGTCGATAGGAAGGAAGAACGCAGTGCA   352
           |||  ||| |||  |||:::|||         :::|||  |||:::|||
Db    125  ArgIleGlnGlnGluSerGlyCysLysSerIleGlnIle--AlaProAspSerGlyGlyLeu   143

Qy    353  GCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGCTGTCCTCCGCTTGTAAGATG       412
           ||||||||| :::|||||||||||:::
Db    144  ProGluArgSerCys***LeuThrGlyThrProGluSerValGlnSerAlaLysArgLeu    163

Qy    413  ATCTTGAGATTATGCATAAA------GAGGCTAAGGACACCCAAAACGGCTGAC------   460
           ::::|||  :::|||||||                  :::
Db    164  LeuAspGlnIleValGlnIleLysArgProAlaProGlyPheHisGlyAspGlyPro      183

Qy    461  ---GAGGTTCCCGAGAGATCCTGGCCCATAATAACTTTGTAGGGGTCTCATTGGCAAG    517
              :::||||||    :::|||  :::                        :::|||||||
Db    184  GlyAsnAlaValGlnIleGluIleMetIleProAlaSerLysSerAlaGlyLeuValIleGlyLys   203

Qy    518  GAAGGACCGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTCG   577
           :::|||  |||||||:::|||       :::                    ::: :::
Db    204  GlyGlyGluThrIleLysGlnIleLeuGlnGluAlaArgGlyAlaLysIleLysValMet---   221

Qy    578  TTGCAAGACCTTACCCTTTACAACCCTGAGAGGACC---------ATCACTGTG         622
           :::|||  |||:::|||                  |||         ::: |||
Db    222  IleGlnAsp---------GlyProGlnAsnThrGlyAlaAspLysProLeuArgIle       237

Qy    623  AAGGGGCCATCAGAATTGTTGCAGGGCCGAGCAGCAAATAATGAAGAAAGTTCGGGAG     682
           :::||||||:::|||:::|||     :::|||  |||:::|||
Db    238  ThrGlyAspProTyrArgLysValGlnLnAlaLysGluMetGluLeuValLeuGluLeuArgAsp   257

Qy    683  ---------GCCTATGAGAATGATGTGGCTGCCATGAATGAGCTCTCCACCTGATCCCT   727
           |||:::
Db    258  GlnGlyGlyPheArgGluValArgArgAsnGluTyrGly---                      269

Qy    728  GGCCTGAACCTGGCTGCTGTGAGGTCTTTTCCCAGTTCATCCAGCGCAGTCCCGCGCCT    787
           |||                                                          :::
Db    269  ---                                                             269

Qy    788  CCCAGCAGCGTTACTGGGCGTGCTCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAG    847
           |||
Db    270  ---SerArgIleGlyGly---                                           276

Qy    848  ATGGTGCAGGTGTTTATCCCGCCAGGCAGTGGCGCCATCATCGGCAAGAAGGGGCAG      907
           :::|||||||||:::|||:::|||||||:||||||:::||||||||||||||
Db    277  GlyIleAspValProIleProArgPheAlaValGlyIleIleValIleGlyAlaGlyAsnGlyGlyGlu   296

Qy    908  CACATCAAACAGCTCTCCCGGTTGCCAGCGCCTCCATCAAGATTGCCACCCCGAA---    964
           :::                                              :::
Db    297  MetIleLysLysIleGlnAsnAspAlaGlyValArgIleGlnPheLysProAspGly     316

Qy    965  ---ACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAG--GCCCAA   1018
                                                              |||
Db    317  ---ArgIleAlaGlnIleThrGlyProProAspArgCysGln                      333

Qy    1019 TTCAAGGCTCAG                                                     1030
           :::
Db    334  HisAlaAlaGluIleIleThrAspLeuArgSerValGlnAlaAlaAsnProGlyGly      353

Qy    1031 ------GGAAGAATCTATGGCAAACTCAAGGAGGAGAGAACTTCTTGGTCCCAAG         1078
                 |||||||
Db    354  ProGlyProGlyGlyArgGlyArgGlyArgGlyLysAsnTrpAsnMetGlyProPro      373

Qy    1079 GAGGAAGTGAAGCTCGAGACCCACATACGTGCCAGCATCAGCAGCTCGCCGGGTCATT    1138
           :::|||     :::|||:::|||
Db    374  GlyGlyLeuGln---GluPheAsnPheIleValProThrGlyLysThrGlyLeuIleIle   392

Qy    1139 GGCAAAGGTGGAAAAACGGTGAACAGTTGCAGAATTGACACGGACGCTGAGGTGGTAGTA  1198
           |||||||||||||
Db    393  GlyLysGlyGlyGluThrValLysIleSerGlnSerGlyAlaArgIleGluLeu         412

Qy    1199 CCAAGAGACCAGACCCCTGATGAGACGACCAGGTC--ATCGTGAAAATCATCGGACAT    1255
           :::|||  |||:::|||:::                  :::|||||||
Db    413  GlnArgAsnProProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr   432

Qy    1256 ---TTCTATGCCAGTCAGATGGCTCAACGGAAGAT------CCGAGACAT            1296
                                                   :::|||
Db    433  ProGlnGlnIleAspTyrArgAlaArgGlnLeuIleGluGluLysIleGlyValAsn      452

Qy    1297 CCT---GGCCCAGGTTAAGCAGCATCAGAAGGGACAGCAGAGTAACCAGGCCCAGGCACG  1353
           |||   |||                                              :::
Db    453  ProLeuGlyProProValProHisGlyProHisGlyVal------ProGlyPro--HisG  470

Qy    1354 GA------GGAAGTGACCAGCCCCTC-----CCTGTCCCTTNGA                 1386
           |||     :::|||
Db    470  lyProProGlyProGlyProThrProMetGlyProTyrAsnProAlaProTyrA        490

Qy    1387 GTCCAGGA---                                                     1394
           :::
Db    490  snProGlyProProGlyProProAlaProHisGlyProProAlaProTyrAlaProGlnGly7  510

Qy    1395 ---CAACAGCGGCAGAAATGCAGAGTGTGCTCTCCCGG                         1431
              ||||||:::|||
Db    510  rpGlyAsnAlaTyrProHisTrpGlnGlnAlaProProAsp---                   526

Qy    1432 CAGGCCTGAGAATGAGTGGGAATCCGGGACACNTGGGCCGGGGTCGTAGATCAGGTTTGCC  1491
           :::                                              :::
Db    526  laLysLeuAlaGlyThrAspProAsnSerAla-AlaTrpProAlaAlaTyr------      542

Qy    1492 CACTTGATTGAGAAAGATGTTCCAGTGAGGACCCTGATCTNTCAGCCCAAACACCCAC    1551
           :::                                              :::
Db    543  HisTyrTyrGlyGlnGlnAla---                                        555

Qy    1552 CCAATTGGC---CCAACACTGTNTGCCCCTCGGGTGTCAGAAATTNTAGCGCAAGGCAC   1608
           :::|||     :::
Db    556  ProAlaGlyAlaProThrThrThrGlnThrAsnGlyLnGlyAsp---                570
```

us-09-270-437d-5_1.rai

```
Qy    1609 TTTTAAACGTGGATTGTTTAAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACC 1668
                 :::   ::::|||  ||||  :::||||||:::|||
Db     571 ----------------GlnGlnAsnProAlaProAlaGlyGlnValAspTyrThr 583

Qy    1669 ----------TCAGTGGGAAGAAGAAATAAAATTTCCTTCAGGT 1701
                 :::||||:::   :::    |||:::|||
Db     584 LysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaValProAlaProThrGly 602
```

Search completed: July 13, 2004, 12:14:43
Job time : 47 secs

q

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:     July 13, 2004, 17:54:49 ; Search time 134 Seconds
                                      (without alignments)
                            7073.553 Million cell updates/sec

Title:          US-09-270-437D-5
Perfect score:  1708
Sequence:       1 aggacgctgccgcaccgcc........atttcctcaggtttttaaaa 1708

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       682709 seqs, 27747446 residues

Total number of hits satisfying chosen parameters:    1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
     1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
     2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
     3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
     4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
     5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
     6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	1703	99.7	1708	3	US-09-061-709-5	Sequence 5, Appli
2	1703	99.7	1708	4	US-09-899-651-5	Sequence 5, Appli
3	1669.2	97.7	1946	4	US-09-061-709-7	Sequence 7, Appli
4	1669.2	97.7	1946	4	US-09-899-651-7	Sequence 7, Appli
5	1165	68.2	2224	3	US-09-261-855-1	Sequence 1, Appli
6	582.2	34.1	1740	4	US-09-643-597-347	Sequence 347, App
7	582.2	34.1	1740	4	US-09-542-615A-347	Sequence 347, App
8	582.2	34.1	1740	4	US-09-606-421B-347	Sequence 347, App
9	580.6	34.0	4159	3	US-09-061-709-4	Sequence 4, Appli
10	580.6	34.0	4159	4	US-09-899-651-4	Sequence 4, Appli
11	580.6	34.0	4181	4	US-09-643-597-175	Sequence 175, App
12	580.6	34.0	4181	4	US-09-480-884A-175	Sequence 175, App
13	580.6	34.0	4181	4	US-09-542-615A-175	Sequence 175, App
14	580.6	34.0	4181	4	US-09-606-421B-175	Sequence 175, App
15	504	29.5	3283	3	US-09-061-709-8	Sequence 8, Appli
16	504	29.5	3283	4	US-09-899-651-8	Sequence 8, Appli
17	489.2	28.6	3412	3	US-09-061-709-6	Sequence 6, Appli
18	489.2	28.6	3412	3	US-09-899-651-6	Sequence 6, Appli
19 c	53.6	3.1	7218	1	US-08-232-463-14	Sequence 14, Appl
20	45.6	2.7	598	4	US-09-669-751-39	Sequence 39, Appl
21	40.6	2.4	2850	1	US-08-503-172-5	Sequence 5, Appli
22	40.6	2.4	2850	2	US-09-135-211-5	Sequence 5, Appli
23	39.6	2.3	1982	4	US-09-016-434-1067	Sequence 1067, Ap
24	39.6	2.3	1982	4	US-09-825-497A-40	Sequence 40, Appl
25	39.6	2.3	1987	4	US-09-825-497A-39	Sequence 39, Appl
26	39.6	2.3	1987	4	US-09-517-779-1	Sequence 1, Appli
27	37.6	2.2	1896	4	US-09-343-011B-4	Sequence 4, Appli
28	37	2.2	2608	4	US-09-154-750A-75	Sequence 75, Appl
29	37	2.2	2608	4	US-09-665-479A-11	Sequence 11, Appl
30 c	36.6	2.1	364	4	US-09-621-976-17202	Sequence 17202, A
31	36	2.1	555	4	US-09-252-991A-6251	Sequence 6251, Ap
32	36	2.1	1644	4	US-09-252-991A-6420	Sequence 6420, Ap
33	36	2.1	2088	4	US-09-252-991A-6339	Sequence 6339, Ap
34 c	36	2.1	2202	4	US-09-252-991A-6131	Sequence 6131, Ap
35	35.8	2.1	2463	4	US-09-252-991A-6775	Sequence 6775, Ap
36 c	35.6	2.1	9862	4	US-09-691-861A-3	Sequence 3, Appli
37 c	35.2	2.1	505	4	US-09-621-976-15639	Sequence 15639, A
38	35.2	2.1	2517	4	US-09-894-998A-38	Sequence 38, Appl
39	35	2.0	2655	4	US-09-963-137-139	Sequence 139, App
40	35	2.0	2655	4	US-09-963-137-183	Sequence 183, App
41	34.6	2.0	832	4	US-09-621-976-2813	Sequence 2813, Ap
42	34.6	2.0	1926	4	US-09-249-585A-2	Sequence 2, Appli
43	34.6	2.0	1926	4	US-09-410-399-3	Sequence 3, Appli
44	34.6	2.0	2319	3	US-09-070-356-1	Sequence 1, Appli
45	34.6	2.0	2580	3	US-09-050-863-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-061-709-5
; Sequence 5, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LJD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-5

Query Match             99.7%;  Score 1703;  DB 3;  Length 1708;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1708;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy         1 AGGGACGCTGCGCACCGCCCCAGTTTACCCGGGAGCCATCATGAAGCTGAATGGCCA 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 AGGGACGCTGCGCACCGCCCCAGTTTACCCGGGAGCCATCATGAAGCTGAATGGCCA 60

Qy        61 CCAGTTGGAGAACCATGCCCTGAAGGTCTCTACAATCCCGATGAGCAGATAGCACAGGG 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 CCAGTTGGAGAACCATGCCCTGAAGGTCTCTACAATCCCGATGAGCAGATAGCACAGGG 120

Qy       121 ACCTGAGAATGGGCGGCGGGCGCCGGCCCAAGCAGCAGCAAGTGACATCCCCTTCGGCTCCTCACC 180
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 ACCTGAGAATGGGCGGCGGGCGCCGGCCCAAGCAGCAGCAAGTGACATCCCCTTCGGCTCCTCACC 180

Qy       181 TGTGGCAGCGCGGGGCGCCGGGGCCCCAGCCAAGCAGCAGCAAGTGACATCCCCTTCGGCTCCTCTGGT 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 TGTGGCAGCGCGGGGCGCCGGGGCCCCAGCCAAGCAGCAGCAAGTGACATCCCCTTCGGCTCCTCTGGT 240

Qy       241 GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGAGGGGCCACCATCCCAACATCAC 300
             |||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGAGGGGCCACCATCCCAACATCAC 300

Qy       301 AAAACAGCACCCAGTCCAAGATAGACGTGCATAGGAGGAGAACGCAGGTGCAGTGCAGTGAAAA 360
```

```
Db   1381  CTTNGAGTCCAGGACAACAACGGGGCAGAAATCGAGAGTGTGCTCTCCCGGCAGGCCTGA  1440
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1441  GAATGAGTGGGAATCCGGGACACNTGGGCGGCCGGGCTGTGTAGATCAGGTTTGCCCACTTGATT  1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  GAATGAGTGGGAATCCGGGACACNTGGGCGGCCGGGCTGTGTAGATCAGGTTTGCCCACTTGATT  1500
Qy   1501  GAGAAAGATGTTCAGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCACCCAATTGGC  1560
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  GAGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCACCCAATTGGC  1560
Qy   1561  CCAACACTGTNTGCCCCTCCGGGTGTCAGAAATNTAGGCAAGGCACTTTTAAACGTGTG  1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  CCAACACTGTNTGCCCCTCCGGGTGTCAGAAATNTAGGCAAGGCACTTTTAAACGTGTG  1620
Qy   1621  ATTGTTTAAAGAAGCTCTCAGGCCCCAAGAGAGGGTGATCACACCTCAGTGGGAAGA  1680
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  ATTGTTTAAAGAAGCTCTCAGGCCCCAAGAGAGGGTGATCACACCTCAGTGGGAAGA  1680
Qy   1681  AAAATAAAAATTTCCTTCAGTTTTAAAA  1708
            ||||||||||||||||||||||||||||
Db   1681  AAAATAAAAATTTCCTTCAGTTTTAAAA  1708
```

RESULT 2
US-09-899-651-5
; Sequence 5, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
;   APPLICANT: Chen, Yao-Tseng
;   APPLICANT: Gure, Ali
;   APPLICANT: Tsang, Solam
;   APPLICANT: Stockert, Elisabeth
;   APPLICANT: Jager, Elke
;   APPLICANT: Knuth, Alexander
;   APPLICANT: Old, Lloyd J.
;   TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
;   TITLE OF INVENTION: Associated
;   TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
;   FILE REFERENCE: LUD 5538
;   CURRENT APPLICATION NUMBER: US/09/899,651
;   CURRENT FILING DATE: 2001-07-06
;   PRIOR APPLICATION NUMBER: US/09/061,709
;   PRIOR FILING DATE: 1998-04-17
;   NUMBER OF SEQ ID NOS: 8
;   SEQ ID NO 5
;   LENGTH: 1708
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-899-651-5

Query Match   99.7%; Score 1703; DB 4; Length 1708;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1  AGGGACGCTGCCGCACCGGCCCCAGTTTACCCCGGGGAGCCATCATGAAGCTGAATGGCCA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  AGGGACGCTGCCGCACCGGCCCCAGTTTACCCCGGGGAGCCATCATGAAGCTGAATGGCCA  60
Qy    61  CCAGTTGGAGAACCATGCCGCCAGTCTCCTACATCCCGGGGTCTCTACATCCCGATGAGCAGCAGAGGG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CCAGTTGGAGAACCATGCCGCCAGTCTCCTACATCCCGGGGTCTCTACATCCCGATGAGCAGCAGAGGG  120
Qy   121  ACCTGAGAATGGCGCCGAGGGGGCTTTGGCTCTTGGCTCTCAGCAGCAGCA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  ACCTGAGAATGGCGCCGAGGGGGCTTTGGCTCTTGGCTCTCAGCAGCAGCA  180
Qy   181  TGTGCAGCAGCAGAGTAACCAGGCCCAGCAGAGTGACACATCCCCCTTCGGCTCCTGT  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  TGTGCAGCAGCAGAGTAACCAGGCCCAGCAGAGTGACACATCCCCCTTCGGCTCCTGT  240
Qy   241  GCCCACCCAGTATGTGGGTGCCATTATTGGTGCAAGGAGGGGGCCAACATCAC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  241 GCCCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCAC 300
Qy  301 AAAACAGACCCAGTCCAAGATAGAGAGGAAGGAGAACGCAGGTGCAGCTGCAGCTGAAAA 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 AAAACAGACCCAGTCCAAGATAGAGAGGAAGGAGAACGCAGGTGCAGCTGCAGCTGAAAA 360

Qy  361 AGCCATCAGTGTGCACCCCTGAGGCTGCTCCTCCGCTTGTAAGATGATCTTTGA 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 AGCCATCAGTGTGCACCCCTGAGGCTGCTCCTCCGCTTGTAAGATGATCTTTGA 420

Qy  421 GATTATGCATAAAGAGAGGCTAAGGACACCAAAACGGCTGACGACGTTCCCCTGAAGATCCT 480
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GATTATGCATAAAGAGAGGCTAAGGACACCAAAATCGCTGACGACGTTCCCCTGAAGATCCT 480

Qy  481 GGCCCATATAACTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAGAGGT 540
        |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  481 GGCCCATATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAGAGGT 540

Qy  541 AGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTTTACAA 600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db  541 AGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTTTTACAA 600

Qy  601 CCCTGAGAGGACCACCATCACTGTGAAGGGGGCCATCGTGAAGGGCCGAGCAGGA 660
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CCCTGAGAGGACCACCATCACTGTGAAGGGGGCCATCGTGAAGGGCCGAGCAGGA 660

Qy  661 AATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGCTCCACTT 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 AATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGCTCCACCT 720

Qy  721 GATCCCTGGCCTGAACCTGGCTGCGCTGCGTGTAGGTCTTTTCCCAGCTTCATCCAGGCAGTCCC 780
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GATCCCTGCTGCCTGAACCTGGCTGCGCTGCGTGTAGGTCTTTTCCCAGCTTCATCCAGCAGTCCC 780

Qy  781 GCCGCCTCCCAGCAGCGGTTACTGGGGCGTCTGCTGCTCCTTTATGCAGGCTCCCCGA 840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GCCGCCTCCCAGCAGCGGTTACTGGGGCGTCTGCTGCTCCTTTATGCAGGCTCCCCGA 840

Qy  841 GCAGGAGATGGTGCAGGTGTTATCCCCCGCCAGCAGTGGGGCGCCCATCATCCGCCAAGAA 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GCAGGAGATGGTGCAGGTGTTATCCCCCGCCAGCAGTGGGGCGCCCATCATCCGCCAAGAA 900

Qy  901 GGGGCAGCACATCAAAACAGCTCTCCCGGTTGCCAGCGGCCTCCATCAAGAGATTGCACCACC 960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 GGGGCAGCACATCAAAACAGCTCTCCCGGTTGCCAGCGGCCTCCATCAAGAGATTGCACCACC 960

Qy  961 CGAAACACCTGACTCCAAAGTTCGTATGTGTTATCATCACTGACCGCCGCCCAGAGCCCCAATT 1020
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CGAAACACCTGACTCCAAAGTTCGTATGTGTTATCATCACTGACCGCCGCCCAGAGCCCCAATT 1020

Qy 1021 CAAGGCTCAGGGAGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCCAAGGA 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 CAAGGCTCAGGGAGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCCAAGGA 1080

Qy 1081 GGAAGTGAAGCTGAGAGACCCCACCATACGTGTGCCAGCATCAGCAGCTGGCCGGGTCATTGG 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 GGAAGTGAAGCTGAGAGACCCCACCATACGTGTGCCAGCATCAGCAGCTGGCCGGGTCATTGG 1140

Qy 1141 CAAAGGTGAAAAACGGTGAACCAGTTGCAGCGGCAGAATTGACGGCAGCTGAGGTGTGTAGTACC 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CAAAGGTGAAAAACGGTGAACCAGTTGCAGCGGCAGAATTGACGGCAGCTGAGGTGTGTAGTACC 1200

Qy 1201 AAGAGACCAGACCCCTGATGAGACGACAGCAGGTCATCGTGAAAATCATCGGACATTCTA 1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 AAGAGACCAGACCCCTGATGAGACGACAGCAGGTCATCGTGAAAATCATCGGACATTTCTA 1260

Qy 1261 TGCCAGTCAGTGCTCAACGGAAGATCCGAGACATCCGAGAACATCCGAGGCACCGGAGCAGCA 1320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 TGCCAGTCAGTGCTCAACGGAAGATCCGAGACATCCGAGAACATCCGAGGCACCGGAGCAGCA 1320

Qy 1321 TCAGAAGGGACAGAGTAACCAGGCCCAGGCCCACCGGACGAGGAAGTGCAGCCCCTGTCC 1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 TCAGAAGGGACAGAGTAACCAGGCCCAGGCCCACCGGACGAGGAAGTGCAGCCCCTGTCC 1380
```

```
Qy 1381 CTTNGAGTCAAGGAGCACCACAACGGGCGACGAAATCGAGCAGTGTGCTCTCCCGGCAGGCCTGA 1440
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 CTTNGAGTCAAGGAGCACCACAACAACGGGCAGAAATCGAGCAGTGTGCTCTCCCGGCAGGCCTGA 1440

Qy 1441 GAATGAGTGGGAATCCGGGACACNTGGGCCGGCCGGCTGTAGATCAGGTTGCCCACTCAGGATTGATT 1500
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GAATGAGTGGGAATCCGGGACACNTGGGCCGGCCGGCTGTAGATCAGGTTGCCCACTCAGGATTGATT 1500

Qy 1501 GAGAAAGATGTTCCAGTGAGGAACCCTGATCNTCAGCCCCAAACACACCCAATTGGC 1560
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 GAGAAAGATGTTCCAGTGAGGAACCCTGATCNTCAGCCCCAAACACACCCAATTGGC 1560

Qy 1561 CCAACACTGTNTGCCCCTCGGCGGTGTCAGAATTNTAGCGCAAGGCACTTTTAAACGTGG 1620
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 CCAACACTGTNTCCCCTCGGCGGTGTCAGAATTNTAGCGCAAGGCACTTTTAAACGTGG 1620

Qy 1621 ATTGTTTAAAGAAGCTCTCCAGGCCCCACCAAGAGAGGTGGATCACACCTCAGTGGGAAGA 1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 ATTGTTTAAAGAAGCTCTCCAGGCCCCACCAAGAGAGGTGGATCACACCTCAGTGGGAAGA 1680

Qy 1681 AAAATAAAATTCCTTCAGGTTTTAAAA 1708
        |||||||||||||||||||||||||||
Db 1681 AAAATAAAATTCCTTCAGGTTTTAAAA 1708
```

RESULT 3
US-09-061-709-7
; Sequence 7, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-7

Query Match   97.7%; Score 1669.2; DB 3; Length 1946;
Best Local Similarity   99.8%; Pred. No. 0;
Matches 1676; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy   30 CCCGGGGAGCCATCATGATGAAGCTGAATAGCCACCAGTTGGAGACCATGCCCCTGAAGGTCT 89
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  268 CCAGGCAAGCCATCATGATGAAGCTGAATAGCCACCAGTTGGAGAATGGCCACCATGCCCCTGAAGGTCT 327

Qy   90 CCTACATCCCCGATGAGCAGATAGCACAGGACCTCAGGGCTCACCTGTGCGACCCCAAGTCT 149
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  328 CCTACATCCCCGATGAGCAGATAGCACAGGACCTCAGGGCTCACCTGTGCGACCCCAAGTCT 387

Qy  150 GGTCTCCGGGGTCAGCGCCCCAGGAGCGTCAGCCTCGTGGCCCAGCCAGCAGC 209
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  388 GGTCTCCGGGGTCAGCGCCCCAGGAGCGTCAGCCTCGTGGCCCAGCCAGCAGC 447

Qy  210 AGCAAGTGGACATCCCCCTTCGGCTCCTCGGTGCCCACCAGTTGGTGCCATTATTG 269
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  448 AGCAAGTGGACATCCCCCTTCGGCTCCTCGGTGCCCACCAGTTGGTGCCATTATTG 507

Qy  270 GCAAGGAGGGGGCCACCATCCGCAACATCACCAAAAACAGACCCAGTCCAAGATAGACGTGC 329
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  270 GCAAGGAGGGGGCCACCATCCGCAACATCACCAAAAACAGACCCAGTCCAAGATAGACGTGC
```

```
Db    508  GCAAGGAGGGGGCCACCATCCGCAACATCACAAAAACAGACCCAGTCCAAGATAGACGTGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1410  ATCGAGAGTGTGCTCTCCCGGCAGGCCTGAGATGAGTGGAATCCGGGACACNTGGGC 1469
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Qy   1648  ATCGAGAGTGTGCTCTCCCGGCAGGCCTGAGATGAGTGGAATCCGGGACACNTGGGC 1707

Qy    330  ATAGGAAGGAAGCAGCGCAGTGCAGCGCAGCCATCAGTGTGCACTCCACCCCTGAGG 389
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    568  ATAGGAAGGAAGCAGCGCAGTGCAGCGCAGCCATCAGTGTGCACTCCACCCCTGAGG 627

Db   1470  CGGGCTGTAGATCAGGTTGCCCACTTGCCCACTTGATTGAGAAAGATGTTCAGTGAGGAACCCTGA 1529
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1708  CGGGCTGTAGATCAGGTTGCCCACTTGCCCACTTGATTGAGAAAGATGTTCAGTGAGGAACCCTGA 1767

Qy    390  GCTGCTCTCCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 449
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    628  GCTGCTCTCCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 687

Db   1530  TCTNTCAGCCCAAACACCCACCAATTGCCCAACACTGTNTGCCCTTCGGGGTGTCAG 1589
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1768  TCTNTCAGCCCAAACACCCACCAATTGCCCAACACTGTNTGCCCTTCGGGGTGTCAG 1827

Qy    450  AAACGGCTGACGAGGTTCCCTGAAGATCCTGCCCATAATAACTTTGTGTAGGGGTCTCA 509
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    688  AAACGGCTGACGAGGTTCCCTGAAGATCCTGCCCATAATAACTTTGTGTAGGGGTCTCA 747

Db   1590  AAATNTAGGCGAAGCGCACTTTAAAGTGTGGATTGTTTAAAGAAGCTCTCCAGGCCCAC 1649
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1828  AAATNTAGGCGAAGCGCACTTTAAAGTGTGGATTGTTTAAAGAAGCTCTCCAGGCCCAC 1887

Qy    510  TTGGCAAGGAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAATCACCA 569
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    748  TTGGCAAGGAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAATCACCA 807

Qy   1650  CAAGAGGGTGGATCACACCTCAGTGGGAAGAAAATAAAATTCCTTCAGGTTTTTAAAA 1708
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1888  CAAGAGGGTGGATCACACCTCAGTGGGAAGAAAATAAAATTTCCTTCAGGTTTTTAAAA 1946

Qy    570  TCTCCTCGTTGCAAGACCTTACAACCCTGAGAGACCATCACTGTGAAGGGGG 629
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db    808  TCTCCTCGTTGCAAGACCTTACAACCCTGAGAGACCATCACTGTGAAGGGGG 867

Qy    630  CCATCGACGAAATTGTTGCAGGGCGAGCAGGAGAATAAATGAAGAAAGTTCGGGAGGCCTATG 689
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    868  CCATCGACGAAATTGTTGCAGGGCGAGCAGGAGAATAATGAAGAAAGTTCGGGAGGCCTATG 927

Qy    690  AGAATGATGTGGCTGCCATGAGCTCTCACCTGACCTGCTGCTGTAG 749
                ||||||||||||||||||||||||||||||||||||||||||||
Db    928  AGAATGATGTGGCTGCCATGAGCTCTCACCTGACCTGCTGCTGTAG 987

Qy    750  GTCTTTTCCCAGCTTCATCCAGGCAGTCCCCGCTCCCCAGCAGCGTTACTGGGGCTG 809
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    988  GTCTTTTCCCAGCTTCATCCAGGCAGTCCCCGCTCCCCAGCAGCGTTACTGGGGCTG 1047

Qy    810  CTCCCTATAGCTCCTTTATGCAGGCTCCGAGGAGATGGTGCAGTGTGTTTATCCCCG 869
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1048  CTCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAGATGGTGCAGTGTGTTTATCCCCG 1107

Qy    870  CCCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGCAGCACATCAAACAGTCTCCCGGT 929
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1108  CCCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGCAGCACATCAAACAGTCTCCCGGT 1167

Qy    930  TTGCCGCCAGGCCCTCCATCAAGATGCACCACCCGAAACACCTGACTCAAAGTTCGTATGG 989
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1168  TTGCCGCCAGGCCCTCCATCAAGATGCACCACCCGAAACACCTGACTCAAAGTTCGTATGG 1227

Qy    990  TTATCACTGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAAC 1049
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1228  TTATCACTGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAAC 1287

Qy   1050  TCAAGGAGGAGACTTCTTTGGTCCCAAGGAGGAAGTGGAGGACCCACATACGTG 1109
                |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1288  TCAAGGAGGAGACTTCTTTGGTCCCAAGGAGGAAGTGGAGGACCCACATACGTG 1347

Qy   1110  TGGCAGCATCAGCAGCTGCCCGGTCATTGGCAAAGGTGGAAAAACCGTGAACGAGTTGC 1169
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1348  TGGCAGCATCAGCAGCTGCCCGGTCATTGGCAAAGGTGGAAAAACCGTGAACGAGTTGC 1407

Qy   1170  AGAATTTGACGGAGCTGAGGTGGTAGTAGCCAAGGACACCAGAGAAGACC 1229
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1408  AGAATTTGACGGAGCTGAGGTGGTAGTAGCCAAGGACACCAGAGAAGACC 1467

Qy   1230  AGGTCATCGTGAAAATCATCGACATTTCTATGCCAGTTCAGATGGCTCAAGGAAGATCC 1289
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1468  AGGTCATCGTGAAAATCATCGACATTTCTATGCCAGTTCAGATGGCTCAAGGAAGATCC 1527

Qy   1290  GAGACATCCTGGCCCCCAGGTAAGCAGCATCAGAGGGACAGAGTAACCAGGCCCAGG 1349
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1528  GAGACATCCTGGCCCCCAGGTAAGCAGCATCAGAGGGACAGAGTAACCAGGCCCAGG 1587

Qy   1350  CACGGAGGAAGTGACCAGCCCCCTTCCCTTNGAGTCCTTNGAGTCCCTTNGAGTCCAAGACAACAACGGGGCAGAA 1409
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1588  CACGGAGGAAGTGACCAGCCCCCTTCCCTTNGAGTCCCTTNGAGTCCAAGACAACAACGGGGCAGAA 1647
```

RESULT 4
US-09-899-651-7
; Sequence 7, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-7

Query Match   97.7%; Score 1669.2; DB 4; Length 1946;
Best Local Similarity   99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy     30  CCCGGGGAGCCATCATGAAGCTGAATGGTGCCACCAGTTGGAACCATGCCCTGAAGGTCT 89
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268  CCAGGCAAGCCATCATGAAGCTGAATGGTGCCACCAGTTGGAACCATGCCCTGAAGGTCT 327

Qy     90  CCTACATCCCGATGAGCAGCAGATAGCACAGGGACCTGAGCAGGTGCGCCCCGAGATGGCGCCCGGGCCCGAGCAGCCCCAGAGGCCGCCCCCGCGCCCCGGGCCCCGAGCAGCCCCCAGAGGCCGCCCCCGCGCCCGAGCAGCCCCCAGCCCCTTTG 149
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328  CCTACATCCCGATGAGCAGCAGATAGCACAGGGACCTGAGCAGGTGCGCCCCGAGATGGCGCCCGGGCCCGAGCAGCCCCCAGCCCCTTTG 387

Qy    150  GCTCTCTGCGGGGTCAGCCGCCCGCGCCCGGGTCATTGGCCAGTTGCATGGCTCGTGTGCGCCAGTTATGGCAAAC 209
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    388  GCTCTCTGCGGGGTCAGCCGCCCGCGCCCGGGTCATTGGCCAGTTGCATGGCTCGTGTGCGCCAGTTATGGCAAGC 447

Qy    210  AGCAAGTTGGACATCCCCTTCGGCTCCTTGTGTGCGCCACCCAGTATGTGGTGCCCATTATTG 269
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    448  AGCAAGTTGGACATCCCCTTCGGCTCCTTGTGTGCGCCACCCAGTATGTGGTGCCCATTATTG 507

Qy    270  GCAAGGAGGGGGCCACCATCCGCAACATCACAAAAACAGACCCAGTCCAAGATAGACGTGC 329
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    508  GCAAGGAGGGGGCCACCATCCGCAACATCACAAAAACAGACCCAGTCCAAGATAGACGTGC 567
```

```
Qy  1410 ATCGAGAGTGTGCTCTCCCCGGCAGGCCTGAGAATGAGTGGGAATCCGGGACACNTGGGC 1469
         ||||||||||||| ||||||||||||||||||||||||||||||||||||||||| |||||
Db  1648 ATCGAGAGTGTGCTCTCCCCGGCAGGCCTGAGAATGAGTGGGAATCCGGGACACNTGGGC 1707

Qy  1470 CGGGCTGTAGATCAGGTTGCCACTTGGCCCACTTGATTGAGAAAGATGTTCCAGTGAGAACCCTGA 1529
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1708 CGGGCTGTAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGAACCCTGA 1767

Qy  1530 TCTNTCAGCCCCAAACACCCACCCCAATTGGCCCAACACTGNTGCCCTCGGGGTGTCAG 1589
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1768 TCTNTCAGCCCCAAACACCCACCCCAATTGGCCCAACACTGTNTGCCCTCGGGGTGTCAG 1827

Qy  1590 AAATTNTAGGCGAAGGCACTTTTAAACGTGGATTGTTTAAAGAAGTCTCCAGGCCCCAC 1649
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1828 AAATTNTAGGCGAAGGCACTTTTAAACGTGGATTGTTTAAAGAAGTCTCCAGGCCCCAC 1887

Qy  1650 CAAGAGGGTGGATCACACCTCAGTGGAAGAAAATAAAATTTCCTTCAGGTTTTAAA 1708
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1888 CAAGAGGGTGGATCACACCTCAGTGGAAGAAGAAAATAAAATTTCCTTCAGGTTTTTAAA 1946
```

RESULT 5
US-09-261-855-1
; Sequence 1, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-261-855-1

Query Match   68.2%;   Score 1165;   DB 3;   Length 2224;
Best Local Similarity 82.8%;   Pred. No. 0;
Matches 1411;   Conservative   0;   Mismatches 260;   Indels 33;   Gaps 6;

```
Qy    30 CCCGGGGAGCCATCATGAAGTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGTCT 89
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   525 CCAGGCAAGCTATCATGAAGTCATGAAGTAAATGGCCATCAACTGGAGAACCCTGAAGTCT 584

Qy    90 CCTACATCCCGATGAGCAGATAGCAGAGGACCTGAGAATGGCGGCCGAGGGGCTTTG 149
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   585 CCTTACATCCCGATGAGCAGATTAACACAAGGTCCTGAGAATGGCGTCGTGGAGGCTTTG 644

Qy   150 GTCTCTCGGGGTCAGCCCCGCCAGGGGTCACCTGTGCCAGGCCCCAAGCAGCAGC 209
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   645 GGTCTCGGGGCCCGGCCCAGCCCGGCAAGGGTCGCCCGTGGCAGCAGGGCTCAAGCAAGCAGC 704

Qy   210 AGCAAGTGGACATCCCCTTGGACATCCCTCTCCCGCTCCTGGTGGTCACAAACAGACGACC 269
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   705 AGCCAGTGGACATCCCTCTCCCGGCTCCTGGTGGTCCACGACAGTATCATTG 764

Qy   270 GCAAGGAGGGGGCCACCATCCGCAACATCCGAAACATCACAAAACATCACAACAGACACC 329
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   765 GCAAGGAGGGGTGCCACCATCCGCAACATCACAGTGCACTCCCTGAGG 824

Qy   330 ATAGGAAGGAGAACGCAGGTGCAGTGAAAAAGCCATCAGTGTGCACTCCCTGAGG 389
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   825 ATAGGAAGGAGAATGCGGGGCGCTGCGGAAGAGGCCATTCAGCGTGCATTCAACCCCTGAAG 884

Qy   390 GCTGCTCCTCCGCTGTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 449
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   885 GCTGCTCCTCCGCTGTGCAAGATGATCTTGGAGATTATGCACAACAAGGAGGCAAGGACACCA 944

Qy   450 AAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCCATAATAATAACTTTGTAGGGCGTCTCA 509
```

```
Qy   330 ATAGGAAGGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCTGAGG 389
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 ATAGGAAGGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCTGAGG 627

Qy   390 GCTGCTCCTCCGCTGTAAGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 449
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628 GCTGCTCCTCCGCTGTAAGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 687

Qy   450 AAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCCCATAATAATAACTTTGTAGGGCGTCTCA 509
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688 AAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCCCATAATAATAACTTTGTAGGGCGTCTCA 747

Qy   510 TTGGCAAGAAGGACGGAACCTGAAGAAGTAGAGACAAGATACCGAGACAAAAATCACCA 569
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748 TTGGCAAGAAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAAATCACCA 807

Qy   570 TCTCCTCGTTGCAAGAACCTTACAACCCTTACAACCCTGAGGAGGACCATCACTGTGAAGGGGG 629
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   808 TCTCCTCGTTGCAAGAACCTTACAACCCTTACAACCCTGAGGAGGACCATCACTGTGAAGGGGG 867

Qy   630 CCATCGAGAATTGTTGCAGGGGCCGAGCAGCAGGAAATAATGAAGAAAGTTCGGGAGGCCPATG 689
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   868 CCATCGAGAATTGTTGCAGGGGCCGAGCAGCAGGAAATAATGAAGAAAGTTCGGGAGGCCTATG 927

Qy   690 AGAATGATGTGGCTGCCATGAGCTCTCACCTGATCCCTGGCCTGGCCTGCTGTAG 749
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   928 AGAATGATGTGGCTGCCATGAGCTCTCACCTGATCCCGGCCTGGCCTGCTGTAG 987

Qy   750 GTCTTTTCCCAGTTCATCCAGCGCAGTCCCGCCGCCTCCCAGCGCGTTACTGGGGGCTG 809
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   988 GTCTTTTCCCAGTTCATCCAGCGCAGTCCCGCCGCCTCCCAGCGCGTTACTGGGGGCTG 1047

Qy   810 CTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCG 869
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1048 CTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCG 1107

Qy   870 CCCAGGCAGTGGGCGCGCCATCATCGGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGT 929
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1108 CCCAGGCAGTGGGCGCGCCATCATCGGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGT 1167

Qy   930 TTGCCAGCGCCCTCCATCAAGATTGCACCACCCCGAAAAACACCTGACTCCAAAGTTCGTATGG 989
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1168 TTGCCAGCGCCCTCCATCAAGATTGCACCACCCCGAAAAACACCTGACTCCAAAGTTCGTATGG 1227

Qy   990 TTATCATCACTGGACCGGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATGCGCAAAC 1049
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1228 TTATCATCACTGGACCGGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATGCGCAAAC 1287

Qy  1050 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGTGGACCCACATACGTG 1109
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1288 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGTGGACCCACATACGTG 1347

Qy  1110 TGCCAGCATCAGCAGCTCGGCCGGGTCATTGGCAAAGGTGGAAAAACGTGAACGAGTTGC 1169
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1348 TGCCAGCATCAGCAGCTCGGCCGGGTCATTGGCAAAGGTGGAAAAACGTGAACGAGTTGC 1407

Qy  1170 AGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGAACCCCTGATGAGAACGACC 1229
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1408 AGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGAACCCCTGATGAGAACGACC 1467

Qy  1230 AGGTCATCGTGAAAAATCATCGGACACATTTCTATCGCCAGTCAGATGCCTCAACGGAAGATCC 1289
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1468 AGGTCATCGTGAAAAATCATCGGACACATTTCTATCGCCAGTCAGATGGCTCAACGGAAGATCC 1527

Qy  1290 GAGACATCCTGGCCCAGGTTAAGCAGCGCATCAGAGGGCAGAGAGTAACCAGGCCCCAGG 1349
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1528 GAGACATCCTGGCCCAGGTTAAGCAGCGCATCAGAGGGCAGAGAGTAACCAGGCCCCAGG 1587

Qy  1350 CACGGGAAGTGACCAGCCCCTCCGTCCCTTNGAGTCCAGGACAACAACGGGCCAGAA 1409
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1588 CACGGAAGTGACCAGCCCGTCCCCTTNGAGTCCAGGACAACAACGGGCCAGAA 1647
```

RESULT 6
US-09-643-597-347
; Sequence 347, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-643-597-347

Query Match   34.1%;   Score 582.2;   DB 4;   Length 1740;
Best Local Similarity 66.2%;   Pred. No. 2.8e-156;
Matches 858;   Conservative 0;   Mismatches 433;   Indels 6;   Gaps 1;

Query Match   34.1%;   Score 582.2; DB 4; Length 1740;
Best Local Similarity 66.2%; Pred. No. 2.8e-156;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

```
Qy     37 AGCCATCATGAAGCTGAATGGCCACCAGTGAGAGAACCATGCCCTGAAGGTCTCCTACAT 96
          |||||||||| ||||||  ||| ||||||||||||||||||||  |||||||||||||||
Db    402 AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAATTCACCTTGAAAGTAGCCTATAT 461

Qy     97 CCCCGATGAGCAGATAGCACCCAGTCCCCGCTCCCGAACACCCCTTGCGCCGAGGGGGCT 156
          |||| ||||||| |||| ||||||| || ||||||||||||||||||  |||||| ||||
Db    462 CCCTGATGAAACGGCCGCCAGCAAAACCCCTTGCCAGCGGGGCGCCGGGGCT 521

Qy    157 GGGTCAGCCCCGCCAGGGCTCACCGTGTGCAGCAGGGTCTCCAAGGCAGCAGCAGCAAGT 216
          |||||||||||||| |||||| ||||||| |||||||||||||||||||||||| |||||
Db    522 TGGGCAGAGGGGGCTCCTCAAGGCAGGGGTCTCCAAGGATCCGTATCCAAGCAGAAACCATG 581

Qy    217 GGACATCCCCTCGGCTCCTGGTGCCACCCAGTATGTGGGTGCCATTATTGCCAAGGA 276
          |||||||||| |||| |||| |||||| |||  ||||  |||||||| ||||| |||
Db    582 TGATTTGCCTCTGCGCCTGGTTCCCACCCAATTTGTTGAGCCATCATAGGAAAAGA 641

Qy    277 GGGGGCCACCATCCGCAACATCCAACATCACAAAACAGACCCAGTCCAAGACGTGCATAGGAA 336
          |||| ||||||||| ||| |||||||| ||||| |||||||||||||||||||||| ||||
Db    642 AGGTGCCACCATTCGGACACATCGAACATCACAAGACAGACCCAGTCTAAAATGCGATGTCCACCGTAA 701

Qy    337 GGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTCAGTGTGCACTCCACCCCTGAGGGCTGCTC 396
          |||||| |||||||||||||||| |||||||||||| ||||||| ||||||||||||| ||||
Db    702 AGAAAATGCGGGGCGTGCTGCAGAAGTCGATTACTATCCTCCTGAAGGCCACCTC 761

Qy    397 CTCCGCTTGTAAGATGATCTGAGATATGCATAAAGAGCTAAGGACACCAAAACCGC 456
          ||||| ||| |||| || ||||||| ||||| ||||||||||| |||||||||||
Db    762 TGCGGCTTGTAAGTCTATTCTGGAGATTAAGGAAGCTCAAGTATAAATTCAC 821

Qy    457 TGACGAGGTTCCCCTGAGAATCCTGGCCCATATATAACTTGTAGGGGCTCATTGGCAA 516
          || ||||||||||| |||| ||||||||||||||| ||||| || ||||| |||||||
Db    822 AGAAGAGAATCTTAAAAAATTGAGCAAGCAGCACAGACACTAAAATCACGATATCTCC 881

Qy    517 GGAAGGACGGAACCTGACCTGAGGAGTGAGGCTAGAGCAAGATACCGAAGAACGACCA 576
          ||||||| ||||| |||||||||||||||| ||||| ||||||||||  |||||||||
Db    882 AGAAGGAGAAATCTTAAAAAATTGAGCAAGCACAGACACTAAAATCACGATATCTCC 941

Qy    577 GTTGCAAGACCTTACCCTTTACAACCCCAGTCAGATGGCTCAACGGACGGGCCATCGA 636
          |||||||||| |||| ||||||||||  |||||| ||||| ||  |||||||||||||
Db    942 ATTGCAGGAATTGACGCGTGTATAATCCAGAAACGAACGCACTATTACAGTTAAGGCCAATGTTGA 1001

Qy    637 GAATTGTTGCAGGGCCGAGCAGCAGAAATATGAGAAAGTTCGGGAGGCCTATGAGAATGA 696
          ||||                  ||||||||||| |||| |||||||||||| |||||  |
Db   1002 GACATGTGCCAAAGCTGAGGAGATCATGAGGAGATCAGGAGTCTTATGAAAAATGA 1061

Qy    697 TGTGGCTGCCATGA------GCTCTCACCTGACCTCATTTAATTCCTGGATTAAATCTGGATTGTGCTGTAGG 750
          |  || | |||||     ||||| ||  || |||||| |||||| |||||| |||  || ||||||
Db   1062 TATTGCTTCTATGAAATCTTCAAGCACATTTAATTCCTGGATTAAATCTTGGG 1121

Qy    751 TCTTTTCCCAGCTTCATCCAGCGCAGTCCCGGCGTCCCGCCTCCCACCTCAGGGCATCTTGGGGGCTGC 810
          ||  ||||| || |||||||| |  |||||||  ||||||||||||||||||||| |||||| || |
Db   1122 TCTGTTCCCACCACTTCAGGGATGCCACCTCCCCTTCAGCCATGAC 1181

Qy    811 TCCCTATAGCTCCTTTATGCAGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGC 870
          ||||              ||||  |||||||||||||| ||||||||||  |
Db   1182 ...
```

Db 1182 TCCTCCCTACCCGCCAGTTTGAGCAATCAGAAACGGAGACGTGTTCATCTGTTTATCCCAGC 1241
Qy   157 GGGTCAGCCCCGCCCAGGGCTCACCTGTGGGCAGCGGGGGCCCCAGCCAGCCAAGCAGCAGCAAGT 216
Db   522 CCAGGCAGTGGGCGGCCATCATCGGCAAGAAGAGGGCAGCACATCAAACAGCTCTCCCGGTT 930
Qy 1242 TGGGCAGAGGGGGCTCCTCAAGGCAGGGGTCTTCCAGGATCCGTATCCAAGCAGAGAAACCATG 581

Db 1242 TCTATCAGTCGGTGCCATCATCGGCCAGCACATCAAGCAGCTTTCTCGCTT 1301
Qy   217 GGACATCCCCTCGGCTCCTGGTGCCACCAGTATGTGGTGCCATTATTGGCCAAGGA 276

Db   582 TGATTTGCCTCTGCGCTGCCTGGTTCCCACCCAATTTGTTGGAGCCATCATAGGAAAAGA 641
Qy 1182 ...

Db   931 TGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACAACCTGACTCCAAAGTTCGTATGGT 990
Qy   277 GGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAA 336

Db 1302 TGCTGGAGCTTCAATTAAGATTGCTCCAGCACCAGCACCAGGCACCAGCACCAGGATGGT 1361
Qy   642 AGGTGCCACCATTCGGACACATCACCAAACAGACCCCAGTCTAAAATCGATGTCCACGTAA 701

Db 1051 ...
Qy   337 GGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCGAC...GCGGGTGCTC 396

...

RESULT 8
US-09-606-421B-347

; Sequence 347, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-347

Query Match   34.1%; Score 582.2; DB 4; Length 1740;
Best Local Similarity 66.2%; Pred. No. 2.8e-156;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

```
Qy    1231  GGTCATCGTGAAAATCATCGGACATTTCTATGCGCAGTCAGATGGCTCAACGCAAGGAAGATCCG 1290
Db    1602  AGTGTTGTCAAAATAACTGGTCACTTCTATGCTTGCTTGCCAGGTTGCCCAGAGAAAAATTCA 1661

Qy    1291  AGACATCCTGGCCCAGGTTAAGCAGCAGCAGCATCAGAAG 1327
Db    1662  GGAAATTCTGACTTCAGGTAAAGCAGCAGCAACAACAG 1698
```

RESULT 9
US-09-061-709-4
; Sequence 4, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-4

```
Query Match       34.0%;    Score 580.6;  DB 3;    Length 4159;
Best Local Similarity  66.1%;  Pred. No. 1.2e-155;
Matches  857;  Conservative   0;  Mismatches 434;  Indels   6;  Gaps    1;

Qy      37  AGCCATCATGAAGTGAATGGCCACCAGTTGAGAACCATGCCCAGTTGTGAGGTCTTCCTACAT   96
Db     652  AGCACTAGACAAACTGAATCGATTTCAGTTAGAGAATTTCACCTTGAAAGTAGCCTATAT  711

Qy      97  CCCGATGAGCAGATAGCACAGGAGACCTGAGAATGGGCCGCAGACCCCCAGCAGCAGCAAGT  156
Db     712  CCCTGATGAAATGGCCGCCCAGCAACAACCCCTTGCAGCAGCCCCGAGGTCGCCGGGGGCT  771

Qy     157  GGGTCAGCCCCCGCCAGGGGCTCACCTGTGCAGCGCGGGGGCCCAGCCAGCAGCAAGT  216
Db     772  TGGGCAGAGGGGCTCCTCAAGCAGCAGGGGTCTCCAAGCAGCCGTATCCAAGCAGAAACCATG  831

Qy     217  GGACATCCCCCTTCGGCTCCTGGTCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGA  276
Db     832  TGATTTGCCTCGCCGCCTCCTGCGCTCGGTTCCCACCCAATTGTTGGAGCCATCATAGGAAAAGA  891

Qy     277  GGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCAAGATAGACGTGCATAGGAA  336
Db     892  AGGTGCCACCATTCGGAACATCGGAAAACAGAAACATCCAGTTCAAAATCGATGTCCACCGTAA  951

Qy     337  GGAGAACGCAGGTGCAGCTGAAAAAGCCATCGAAAAAGCCATCGTGTGCACTCCCCCCTGCTC  396
Db     952  AGAAAATGCGGGGCTGCTGAGAGTGCTGATTACTATCCTCTACTCCTGAAGGCACCTC  1011

Qy     397  CTCCGCTTGTAAGATGATCATCTTGAAGATTATGCATAAAGAAGGCTAAGGACACCAAAAACGC  456
Db    1012  TGCGGCTTGTAAGTCTATTCTGAGATTATGCATAAGCAAGATCAAGATATAAAAATTCAC  1071

Qy     457  TGACGAGGTTCCCCTGAAGATCCTGCCCATAATAACTTGTTAGGGCGTCTCATTGGCAA  516
Db    1072  AGAAGAGATCCCCCTTGAAGATTTTGAAGATCCCTGAAGGTCTCATAAAACTTTGTTGGACGTCTTATTGGTAA  1131

Qy     517  GGAAGGACCGAACTTGAGAAGGTAGAGCAAGATACCGGACAAGAAATCACCATCTCCTC  576
Db    1132  AGAAGGAAGAAATCTTAAAAAAATTGAGCAAGAGAACAGACACTAAAAATCACGATATCTCC  1191
```

```
Qy     577  GTTGCAAGACCTTACCCTTTACACACCTGAGAGGACCATCACTGTGTGAAGGGGCCATCGA  636
Db    1192  ATTGCAGGAATTGACGTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGA  1251

Qy     637  GAATTGTTGCAGGGCCGAGCAGAGAAATAATGAGCAAAGTTCGGGAGAAAGTTCGGGAGAATGA  696
Db    1252  GACATGTTGCCAAAGCTGAGGAGCAGATCATGAGGAAAAATCAGGGAGTCTTATGAAAATGA  1311

Qy     697  TGTGGCTGCCATGA------GCTCTCACCTGATCCCTGGCTGCCTGGCTGCTGTCTGTAGG  750
Db    1312  TATTGCTTCTATCGAATCTTCAAGCACATTTAATTCCTGGATTAAATCTGAACCGCCTTCGG  1371

Qy     751  TCTTTTCCCAGCTTCATCCAGCCGCAGTCCGCGCCTCCCAGCGAGCGTTACTGGGGCTGC  810
Db    1372  TCTGTTCCCACCCACTTCAGGGATGCCACTCTCCACCTTCAGGGCCCCCTTCAGCCATGAC  1431

Qy     811  TCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAGAGTGTGCAGGTGTTTATCCCCGC  870
Db    1432  TCCTCCCTACCCCGCAGTTTGAGCAATCAGAAACGGAGGACTGTTCATCAGTTTATCCCAGC  1491

Qy     871  CCAGGCAGTGGCGGCCATCATCGGCAAGAAGGGGCAGCACACATCAAACAAGCTCTCCCGGTT  930
Db    1492  TCTATCAGTCGGTGCCATCATCGGCAACATCGGCAGCAGGGCCACATCCAAGCACAGCTTTCTCGCTT  1551

Qy     931  TGCCAGCGCCTCATCAAGATTGCACCGGAAACACCTGACTCCAAAGTTCGTATGGT  990
Db    1552  TGGTGGAGCTTCAATTAAGATTGCTCCAGGCACCAGCTCTAAAGTGACTGAGGATGGT  1611

Qy     991  TATCATCACTGACCGGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT  1050
Db    1612  GATTATCACTGGACACCAGAGGCTCAGTTCAAGGCTCAGGGAGAGAATTTATGGAAAAAT  1671

Qy    1051  CAGGGAGGAGAGAACTCTTGGTCCCAAGGAGGAGAAGTGAAGCTGGAGACCCACATACGTGT  1110
Db    1672  TAAAGAAGAAAACTTTGTTAGTCCTAAAGAAGAAGAAGGTGAAACTTGAAGCTCATATCAGAT  1731

Qy    1111  GCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGTGGAAAAACGGTGAACCAGTTGCA  1170
Db    1732  GCCATCCTTGCTGCTGGTGGCAGAGTTATTGGAAAGAGGCAAAACGGTGAATGAACTTCA  1791

Qy    1171  GAATTTGACGGCAGCTGAGGTGGTAGTACCAGAGACCCCTGATGAAACGACCA  1230
Db    1792  GAATTTGTCAAGTGCAGAAGTTGTTGTCCCTCGACCACCTGATGAGAATGACCA  1851

Qy    1231  GGTCATCGTGAAAATCATCGGACATTTCTATGCGCAGTCAGATGGCTCAACGCAAGGAAGATCCG  1290
Db    1852  AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTTGTTGCCAGGTTGCCCAGAGAAAAATTCA  1911

Qy    1291  AGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAG  1327
Db    1912  GGAAATTCTGACTTCAGGTAAAGCAGCCAACAACAG  1948
```

RESULT 10
US-09-899-651-4
; Sequence 4, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709

```
;  PRIOR FILING DATE: 1998-04-17
;  NUMBER OF SEQ ID NOS: 8
;  SEQ ID NO 4
;  LENGTH: 4159
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
US-09-899-651-4

Query Match         34.0%;  Score 580.6;  DB 4;  Length 4159;
Best Local Similarity  66.1%;  Pred. No. 1.2e-155;
Matches  857;  Conservative  0;  Mismatches  434;  Indels  6;  Gaps  1;

Qy    37 AGCCATCATGAAGCTGAATGGCCACCAGTTGAGAACCATGCCCTGAAGGTCTCCTACAT  96
         ||||||  ||||| || |||||||   |||||||  ||||| ||||||||||||| |||
Db   652 AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAATTTCACCTTGAAAGTAGCCTATAT  711

Qy    97 CCCCGATGAGCAGATAGCACAGGGACCTGAGAGTGGGCGCCGAGGGGGGCTTTGGCTCTCG  156
         ||| |||||||||||  ||||||    |||||||  || ||||||||||||| ||||||||
Db   712 CCCTGATGAAATGCCGCCCAGCAAACCCCTGCAGCAGCGCGGGGCCCCGAGGTCGCCGGGGGCT  771

Qy   157 GGGTCAGCCCCGCCAGGGGCTCACCTGTGCAGCGGGCCCAGCCAAGCAGCAGCAAGT  216
         ||||  ||| ||  |  | |||| ||||| ||||| ||||||| |||||| |||||||
Db   772 TGGGCAGAGGGGGTCCTCAAGGCAGGGGTCTCCAGGATCCGTATCCAAGCAGAAACCATG  831

Qy   217 GGACATCCCCTTCGGCTTGGTGCCACCCAGTATGTGGGTGCCATTATTGGCCAGGA  276
         |||||| |  ||||| || ||||||| || | |  | |||||||||||| ||||||
Db   832 TGATTTGCCTCTGCGCGTTCCTGCTGTCCCACCCAATTTGTTGGAGCCATCATAGGAGAAAGA  891

Qy   277 GGGGGCCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAA  336
         | || ||||||| |||| ||||| ||| ||| | |||  | |||| |  ||||| |||| |
Db   892 AGGTGCCACCATTCGGAACATCACCACCAAACAGACAGACCCCAGTCCAAGTGTCCACCGTAA  951

Qy   337 GGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGCACTCACCCCTGACGGGCTGCTC  396
         ||||| | || |  | | ||| | || |  ||| || ||| || ||| |||||||||
Db   952 AGAAAATGCGGGGGCTGCTGAGAGTCGATGAGATCGA TATCCTCTACTCCTGAGGCCACTC  1011

Qy   397 CTCCGGTCGTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGCACACCAAAACGGC  456
         |||  ||| |||||| | |  | |||||  | || ||| ||| ||||  | ||  |||||||
Db  1012 TGCGGCTTGTAAGTCTATTCTGGAGATTTATGCGAGATTCAAGATATAAAAATTCAC  1071

Qy   457 TGACGAGGTTCCCCTGAAGATCCTGGCCCTCCATTGTGAGGGCGTCTCATTGGCAA  516
         |||||   ||| || ||| ||| | || || | | |||||  || |  ||||| ||
Db  1072 AGAAGAGATCCCCTGAGAGATTTTTAGCTCATAATTAACTTGTTGGACGTCTTATTGGTAA  1131

Qy   517 GGAAGGACGGAACCTGAGAGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTC  576
         |||| | | |||  | |||| ||| || || | | || |||| || | |  ||| | ||
Db  1132 AGAAGGAGAAATCTTAAAAAATTGAGCAAGACAACAGAGACAAATCTAAAATCACGATATCTCC  1191

Qy   577 GTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA  636
         ||| ||||    |  |||  ||| | ||||| || ||||  ||| ||||| |||||| ||
Db  1192 ATTGCAGGAATTGACGCGTGTATATCCGACGAACCGACTATTACAGTTAAAGGCAATGTTGA  1251

Qy   637 GAATTGTTGCAGGGCCGAGCGAGCAGGAAAATAATGAAGAAGTTCGGGAGGCCTATGAGAATGA  696
         || | ||| ||| ||| |||  |||| || | || |||| |  | ||| |||| | |||||| |
Db  1252 GACATGTGCCAAAGCTGGAGGAGATCCATGGAGGAGAAAATCAGGGAGTCTTATGAAAAATGA  1311

Qy   697 TGTGGCTGCCATGA-----GCTCTCACCTGACCCCTGGCGTGCTGTAGG  750
         || | ||||| ||     ||||| || || || ||||||  ||| | ||
Db  1312 TATTGCTTCTATGAATCTTCATCGAGCACATTTAATTCCTGGATTAAATCTGACACGCCTTGGG  1371

Qy   751 TCTTTTCCCAGCTTCATCCAGCGGCAGTCCGGCGCTTCCCACCTACGTGGGGCTGC  810
         ||| |||| ||| | ||||| | ||| ||  || ||| |||| |  |||||||||
Db  1372 TCTGTTCCCACCCACTTCAGGATGCCACCTCCCACCTCAGGCGGCCCCCTTCAGCCCATGAC  1431

Qy   811 TCCCTATAGCTCCTTTATGCAGCTCCCGAGCAGGATGGTGCAGTGTTGTTTATCCCCGC  870
         ||  | || |||  |  | |||||  |||| |  | ||| |  ||| | |||| ||||| |
Db  1432 TCCTCCCTACCCGCAGTTTGAGCAATCAGCAAACGGAGACTGTTCATCAGTTTATCCCCAGC  1491

Qy   871 CCAGGCAGTGGGGCCCATCATGCCCGAGTTGAAGGAAGAGGGCAGCAATCAAGCAGGCTC  930
         |||||| || |||| ||||||  || || || |||| || || || ||||||| ||| ||
Db  1492 TCTATCAGTCGGTGCCATCATCATCCGCAGACCGCAGCAGGGCCAGCAATCACACTCAAGCAGCTTTTCTGCCTT  1551
```

```
RESULT 11
US-09-643-597-175
;  Sequence 175, Application US/09643597
;  Patent No. 6426072
;  GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER:  US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
;  SEQ ID NO 175
;  LENGTH: 4181
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY:  unsure
;  LOCATION:  (3347)
;  OTHER INFORMATION:  n=A,T,C or G
;  NAME/KEY:  unsure
;  LOCATION:  (3502)
;  OTHER INFORMATION:  n=A,T,C or G
;  NAME/KEY:  unsure
;  LOCATION:  (3506)
;  OTHER INFORMATION:  n=A,T,C or G
;  NAME/KEY:  unsure
;  LOCATION:  (3520)
;  OTHER INFORMATION:  n=A,T,C or G
;  NAME/KEY:  unsure
;  LOCATION:  (3538)
;  OTHER INFORMATION:  n=A,T,C or G
;  NAME/KEY:  unsure
```

```
Qy   931 TGCCAGCGCCTCATCAAGATTGACCACCGAAACACCTGACTCCAAAGTTCGTATGGT  990
         ||| ||| ||||| ||||  | ||| | |||||| |||| |||||||||||  | |||
Db  1552 TGCTGGAGCTTCAATTAAGATTGCTCCAGCGAGCCCAATTCAAGGCTCAGGGAGCTCATGGCAAACT  1611

Qy   991 TATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGTTCAGGGAAGAATCTATGCAAACT  1050
         || || |||| |  | ||||| | |  | ||||| || | ||| ||| |||| ||  |||| |
Db  1612 GATTATCACTGGACCACCACGGCTCAGTTCAAGGTCAGGGAAGAATTTATGGAAAAAT  1671

Qy  1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGGAAGTGAGGAGTGAGGACCCACATACGTGT  1110
         ||| ||| ||| |||||  | |  ||||| ||| | |||| ||  | ||| | ||| |||| ||
Db  1672 TAAAGAAGAAACTTTGTTAGTCCTAAAGAAGAGGTGAAACTTGAAGCTCATATCAGAGT  1731

Qy  1111 GCCAGGCATCAGCAGCTGGCCGCGGTCATTGGCAAAGGTGAAAAACGGTGAACGAGTTGCA  1170
         || |  ||||| |||  | | |   |||  |||  ||| |||| | | || ||||| ||
Db  1732 GCCATCCTTTGCTGCTGCGAGCCGCAGAGTTATTGGAAAGGCAAAACGGTGATGATGAACTTCA  1791

Qy  1171 GAATTTGACGGCAGTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGAACGACCA  1230
         ||||| | ||| || | ||  |  ||| | | || |   | || |  ||| |||  |||
Db  1792 GAATTTGTCAAGTGCAGAAGTTGTTGTCCCTGTGACCAGACACCTGATGGCTCACGGAATGACCA  1851

Qy  1231 GGTCATCGTGAAAATCATCGACATTTCTATGCCAGTCAGATGGCTCAACGGAGATCCG  1290
         |||  ||| ||||| |||  ||  || || || |||  |||| ||| ||| ||| | |
Db  1852 AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTCTTGCCAGTTGCCCAGAGAAAAATTCA  1911

Qy  1291 AGACATCCTGCCCAGGTTAAGCAGCAGCAGCAATCAGAAG  1327
         ||| || ||||| ||||  |||| | |||| || | |||
Db  1912 GGAAAATCTGACTCAGGTAAAAGCAGCACCAACAACACAG  1948
```

```
;    LOCATION: (3549)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (3646)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (3940)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (3968)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (3974)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (4036)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (4056)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (4062)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (4080)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (4088)
;    OTHER INFORMATION: n=A,T,C or G
;    NAME/KEY: unsure
;    LOCATION: (4115)
;    OTHER INFORMATION: n=A,T,C or G
;  US-09-643-597-175

Query Match         34.0%;  Score 580.6;  DB 4;  Length 4181;
Best Local Similarity   66.1%;  Pred. No. 1.2e-155;
Matches  857;  Conservative   0;  Mismatches  434;  Indels  6;  Gaps  1;

Qy     37 AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT 96
          ||||||||| |||||||||| |||||||| | |||||||||||||||||||||||||||
Db    652 AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAAGTAGCCTATAT 711

Qy     97 CCCGATGAGCAGATAGCAGAGGGACCTGAGAATGGGCGCCGAGGGGCGTTTGCGTCTCG 156
          || ||| |||||||||||||||||||
Db    712 CCCTGATGAAATGGCCGCCCAGCAAAACCCCTGCAGCGGGGCGCCGGGGGCT 771

Qy    157 GGGTCAGCCCGCGCCAGGGGCTCACCTGTGCAGCGGGGCCCCAGCCAGCAGCAAGT 216
          | | |||||||||
Db    772 TGGGGAGGGGGTCTCCTCAAGGCAGGGGTTCCAGGATCCGTATCCAGAGAAACCATG 831

Qy    217 GGACATCCCCCTTGGCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGA 276
          || |||||||||
Db    832 TGATTTGCCTCTGCGCCTGGTTCCCCACCCAATTTGTTGGAGCCATCATAGGAAAAGA 891

Qy    277 GGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAGATAGACGTGCATAGGAA 336
          ||||||
Db    892 AGGTGCCACCATTGGACATCACCACCACACAGACACCAGTCTAAAATCGATGTCCACCGTAA 951

Qy    337 GGGAGAACGCAGGTGCAGTTGAAAAAGGCCATCAGTGTGCACTCCTCTCTACTCCTGAGGGCTGCTC 396
          ||||| ||||||
Db    952 AGAAAATGCGGGGGCTGCTGAGAGTCGATTACTATCCCTCCTGAAGCCACCTC 1011

Qy    397 CTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAAACGGC 456
          |||| || ||||||
Db   1012 TGCCGCTTGTAAGTCTATTCTGGAGATTATGCATAAGAAGCTCAAGATATAAAATTCAC 1071

Qy    457 TGACCAGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTGTAGGGCGTCTCATTGGCAA 516
          || |||||
Db   1072 AGAAGAGAAATCCCCCTTGAAGAATTTTGACCTTCATAATAACTTTGTTGTTGACCGGTCTTATTGGTAA 1131

Qy    517 GGAAGGACACGAACCTGAAGAAGTAGAGCAAGGCAAGATTACCCGAGACAAAATCACCATCTCCTC 576
          |||| |||||
Db   1132 AGAAGGAAGAAATCTTAAAAAAATTGAGCAAGACACACAAAATCAAAATCACGATATCCC 1191
```

```
Qy    577 GTTGCAAGACCTTACCCTTACACACCCTGAGAGGACCATCACTGTGAAGGGGGCCATCGA 636
          |||||| ||||||||||
Db   1192 ATTGCAGGAATTGACGCTGTATATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGA 1251

Qy    637 GAATTGTTGCAGGGCCGAGCAGGAGAATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGA 696
          |||||||||||||||||
Db   1252 GACATGTGCCAAAGCTGAGGAGGAGATCATGAGGAGAAATCAGGGAGTCTTATGAAAATGA 1311

Qy    697 TGTGGCTGCCATGA------GTTCTCACCTGATCCCTGGCTGAACCTGGCTGCTGTAGG 750
          |||||||||||||    ||| |||||||
Db   1312 TATTGCTTCTATGAATCTTCAAGCACATTTAATTCCTGATTAAATCTGAACGCCTTGGG 1371

Qy    751 TCTTTTCCAGCTTCATCCAGGCCAGTCCCGCGCCTCCCAGCAGCGTTACTGGGCGCTGC 810
          |||||||||||
Db   1372 TCTGTTCCCACCCACTTCAGGGATGCCACCTCCCCACCTCAGGGCCCCCTTCAGCCATGAC 1431

Qy    811 TCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGC 870
          |||||||||
Db   1432 TCCTCCCCTACCCGCAGTTTGAGCAATCAGAAACGAGGGCGCAGCACATCAAACAGCTCTCCCGGTT 1491

Qy    871 CCAGGCAGTGGGCGCCATCATCGGCAAGACAAGGGGCAGCACATCAAACAGCTCTCCCGGTT 930
          |||||||||||
Db   1492 TCTATCAGTCGGTGCCATCATCGCGCAAGCAGGGCCAGCACATCAAGCAGCGTTTCTCGCTT 1551

Qy    931 TGCCAGCCGCCTCCATCAAGATTGCACACCCGAAACAACCTGACTCCCAAAGTTCGTATGT 990
          ||||||||||
Db   1552 TGTGGAGCTTCAATTAAGATTGCTCTCCAGCGGAAGCACCAGAGCTCAGGGATGATGGT 1611

Qy    991 TATCATCACTGACCGGCCAGAGCGCCCAATTCAAGGCTCAGGGAAGAATCTATGCGAAACT 1050
          ||||||||
Db   1612 GATTATCACTGACCACCAGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGGAAAAAT 1671

Qy   1051 CAAGGAGGAGAACTTCTTTGGTCCCCAAGGAGGAAGTGAAGCTCGGAGACCCCACATACGTGT 1110
          ||||||||||
Db   1672 TAAAGAAGAAAACTTTGGTTAGTCCTAAAGAAGAGGTGAAACTTGAAGCTCATATCAGAGT 1731

Qy   1111 GCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAAACGGTGAACAGTTGCA 1170
          |||||||||
Db   1732 GCCATCCTTGCTGCTGGCAGAGTTATTGGAAAAGAGGCAAAACGTGAATGCAACTTCA 1791

Qy   1171 GAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACACCCTGATGAAGAACGACCA 1230
          ||||||||||
Db   1792 GAATTTGTCAAGTGCAGAAGTTGTGTCCTCTGACCAGCACCTGATGAGAATGACCA 1851

Qy   1231 GGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCCAGTCAGTGGCTCAACGAAGATCCG 1290
          |||| ||||||
Db   1852 AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTTGCCACTTCTTGCCCAGAGAAAATTCA 1911

Qy   1291 AGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAG 1327
          |||||||||
Db   1912 GGAAATTCTGACTCAGGTAAAGCAGCCACCAACACAG 1948
```

```
RESULT 12
US-09-480-884A-175
;  Sequence 175, Application US/09480884A
;  Patent No. 6482597
;  GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT: Fan, Liqun
;  APPLICANT: Hosken, Nancy A.
;  APPLICANT: Kalos, Michael D.
;  APPLICANT: Fanger, Gary R.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;  TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C6
;  CURRENT APPLICATION NUMBER: US/09/480,884A
;  CURRENT FILING DATE: 2001-08-27
;  NUMBER OF SEQ ID NOS: 330
;  SOFTWARE: FastSEQ for Windows Version 3.0
;  SEQ ID NO 175
;  LENGTH: 4181
```

```
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc feature
;  LOCATION: (1)...(4181)
;  OTHER INFORMATION: n=A,T,C or G
US-09-480-884A-175

Query Match       34.0%;   Score 580.6;   DB 4;   Length 4181;
Best Local Similarity   66.1%;   Pred. No. 1.2e-155;
Matches  857;   Conservative  0;   Mismatches 434;   Indels  6;   Gaps  1;

Qy      37  AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCTGAAGGTCTCCTACAT    96
                ||||||||||||||| ||||||  |||||| |||||| || || |||||||||||
Db     652  AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAATTTCACCTTGAAAGTAGCCTATAT   711

Qy      97  CCCCGATGAGCAGATAGCACAGGGACCTGAGAATGGCGCCCGAGGGGGGCTTTGGCTCTCG   156
                ||||||   |||| || |||||| ||||||  |||||||  |||||||
Db     712  CCCTGATGAAATGGCCGCCCAGCAAAACCCCTGCAGCAGCCCCCAGCCAAGCAGCAGT   771

Qy     157  GGGTCAGCCCCGCCAGGGGCTCACCTGTGCACCGCGGGCCCCAGCCAAGCAGCAGCAAGT   216
                ||||     || ||| ||||||||
Db     772  TGGGCAGAGGGGTCTCCTCAAGGCAGGGGTCTCCAGGATCCGGATCCAAGCAGAAACCATG   831

Qy     217  GGACATCCCCCTTCGGCTTCCTGGTCCTGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGA   276
                ||||||||||||||||||
Db     832  TGATTTGCCTCTGCGCCTGTGTCCTGTTCCCACCCAATTTGTTGTTGGAGCCATCATAGGAGAAAGA   891

Qy     277  GGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATGACGTGCATAGGAA   336
                ||||||
Db     892  AGTGCCACCATTCGGAACATCACCACAAACAGACCCAGTCCTAAAATCGATGTCCACCGTAA   951

Qy     337  GGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACCCCTGAGGGCTGCTC   396
                ||
Db     952  AGAAAATGCGGGGGCTGCTGAGAAGTCGATTCGATTACTATCCTCTACTCCTGAGGCCACCTC   1011

Qy     397  CTCCGCTTGTAAGTCTATTCTGGAGATTATGCTGGAGGATTATGCATAAAGAGAGCTAAGG   456
                ||||||
Db    1012  TGGCGCTTGTAAGTCTATTCTGGAGATTATGCATAAAGGAAGCTCAAGATATAAAATTCAC   1071

Qy     457  TGACGAGGTTCCCTGAGAATCCTGGCCCAATAATAACTTGTAGGGCGTCTCATTGGCAA   516
                ||||
Db    1072  AGAAGAGATCCCCTTGAGAGATTTTAGTCATCAATAACTTTGTTGGACGTCTTATTGGTAA   1131

Qy     517  GGAAGGACGAACCTGAAGAAGGTAGAGCCAAGTTGAGCAAGAAAATCACCATCTCCTC   576
                ||||||||
Db    1132  AGAAGGAAGAAATCTTAAAAAATTGAGCAAGAAGTTGAGCAAGACACTAAAATCACGATATCTCC   1191

Qy     577  GTTGCAAGACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA   636
                ||
Db    1192  ATTGCAGGAATTGACGCTGTATAATCCAGAACCCACTATTACAGTTAAAGGCAATGTTGA   1251

Qy     637  GAATTGTTGTGCAGGCCGACAGCAGAAATAATGAGAAAGTTCGGGAGGCCTATGAGAATGA   696
                ||||
Db    1252  GACATGTGCCAAAGCTGAGGAGAGATCATGAGGAGAGAAAATACAGGGAGTCTTATGAAAATGA   1311

Qy     697  TGTGGCTGCCATGA------GCTCTCACCTGATCCCTGGCTGCTCGTAGG   750
                ||||||||
Db    1312  TATTGCTTCTATGAATCTTCAAGCACATTAATTCCTGGATTAAATCTGAACGCCTTGGG   1371

Qy     751  TCTTTTCCCAGCTTCATCCAGCGAGCTCCCAGCCAGCCGTTACTGGGGCTGC   810
                ||
Db    1372  TCTGTTCCCACCCACTTCAGGGGTTTCCACTTTATCCCGC   1431

Qy     811  TCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGC   870
                ||
Db    1432  TCCTCCCTACCCGCAGTTTGAGCAATCAGAAACGGAGACTTCATCAGTTTCATCCAGC   1491

Qy     871  CCAGGCAGTGGCGGCCATCATCGGACAGAAGGGGCAGCACATCAAACAGCTTCTCCCGGTT   930
                ||||||
Db    1492  TCTATCAGTCGGTGCCATCATCATCCGCAAGTCAAGCAGGGCCAGCAATCAAGCAGCTTTCTGCTT   1551

Qy     931  TGCCAGCCGCCTCCATCAGATTGCACCACCCCGAAACACCTGACTCCAAGTTCGTATGGT   990
```

```
Db    1552  TGCTGGAGCTTCAATTAAGATTGCTCCAGCGGAGCGAAGCACCAGATGCTAAAGTGAGGATGGT   1611
                 || |  ||||||||||||||||||||||||||||||||||
Qy     991  TATCATCACTGACCCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCATGGCAAACT   1050
                 |||||||||||||||||||||||||||||||||||||
Db    1612  GATTATCACTGACCACCAGGCTTCAGTTCAAGGCTCAAGGGAAGAATTTATGAAAAAT   1671
                 ||||||||
Qy    1051  CAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAGCTGAGCTGAGACCCACATACGTGT   1110
                 |||||||||||||||||||||||||||||||||||||||||||||||||
Db    1672  TAAAGAGAAAACTTTGTTAGTCCTAAAGAAGAGAGGTGAAACTTGAGCTCTATATCAGAGT   1731
                 |
Qy    1111  GCCAGCATCAGCAGCTGGCCGGGTCATTGGCACAAGGTGAAAAGTGGAAAAAACGGTGAACGAGTTGCA   1170
                 |||| ||||||||||||||||||||||||||||
Db    1732  GCCATCCTTTGCTGCTGCGAGAGTTATTGAAAAGGAGGCAAAACGGTCAAATGAACTTCA   1791
                 ||||||||||||||||||||||
Qy    1171  GAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATCGAGAACGACCA   1230
                 || |||||||||||||||||||||||||||||||
Db    1792  GAATTTGTCAAGTGCAGAAGTGTGTTGTCCCTGTCGTGACACACCTGACACCTGACCA   1851
                 |
Qy    1231  GGTCATCGTGAAAATCATCGGACATTTCATGCCAGTCAGTGGCTCAACGGAAGATCCG   1290
                 |||||||||||||||||||||||||||||||
Db    1852  AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTTGCCCAGTTGCCCAGAGAAAAAATTCA   1911
                 |
Qy    1291  AGACATCCTGCCCCAGGTTAAGCACGCAGCATCAGAAG   1327
                 ||||||||||||||||||||||||
Db    1912  GGAAATTCTGACTCAGGTAAAGCACAGCAGCACCAACACAG   1948
```

```
RESULT 13
US-09-542-615A-175
;  Sequence 175, Application US/09542615A
;  Patent No. 6518256
;  GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT: Fan, Liqun
;  APPLICANT: Kalos, Michael D.
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT: Hosken, Nancy A.
;  APPLICANT: Fanger, Gary R.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;  TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C8
;  CURRENT APPLICATION NUMBER: US/09/542,615A
;  CURRENT FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 350
;  SOFTWARE: FastSEQ for Windows Version 3.0
;  SEQ ID NO 175
;  LENGTH: 4181
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: (3347)
;  OTHER INFORMATION: n=A,T,C or G
;  NAME/KEY: unsure
;  LOCATION: (3502)
;  OTHER INFORMATION: n=A,T,C or G
;  NAME/KEY: unsure
;  LOCATION: (3506)
;  OTHER INFORMATION: n=A,T,C or G
;  NAME/KEY: unsure
;  LOCATION: (3520)
;  OTHER INFORMATION: n=A,T,C or G
;  NAME/KEY: unsure
;  LOCATION: (3538)
;  OTHER INFORMATION: n=A,T,C or G
;  NAME/KEY: unsure
;  LOCATION: (3549)
;  OTHER INFORMATION: n=A,T,C or G
;  NAME/KEY: unsure
;  LOCATION: (3646)
;  OTHER INFORMATION: n=A,T,C or G
;  NAME/KEY: unsure
```

```
;   LOCATION: (3940)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3968)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (3974)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4036)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4056)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4062)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4080)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4088)
;   OTHER INFORMATION: n=A,T,C or G
;   NAME/KEY: unsure
;   LOCATION: (4115)
;   OTHER INFORMATION: n=A,T,C or G
US-09-542-615A-175

Query Match      34.0%;  Score 580.6;  DB 4;   Length 4181;
Best Local Similarity   66.1%;  Pred. No. 1.2e-155;
Matches  857;  Conservative   0;  Mismatches  434;  Indels   6;  Gaps   1;

Qy         37 AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT 96
              |||||||||||||||||||||| ||||||||||||||||  ||||||||||||||||
Db        652 AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAATTTCACCTTGAAAGTAGCCTATAT 711

Qy         97 CCCGATGAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGCCCCCCAGCAGCAAGT 156
              |||||| ||| ||||||||||  ||||||||||||||||
Db        712 CCCTGATGAAATGCGCGCCCAGCAAACCCCTGCAGCAGCCCCCGAGTCGCCGGGGGCT 771

Qy        157 GGGTCAGCCCCGCCCAGGGCTCACCTGTGGCCAGCGGGCCCCAGCCAGCAGCAGCAAGT 216
              ||||||  |||| |||
Db        772 TGGGCAGGAGGGGGCTCCTCAAGGCAGGGGTCTCCAGGATCCGTATCCAAGCAGAAACCATG 831

Qy        217 GGACATCCCCTTCGGCTCTCGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGA 276
              |||| ||| |||||
Db        832 TGATTTGCCTCTGCGGCCTGCGCCTGGTGTTCCCACCCAATTGTTGTTGGCCATCATAGGAAAAGA 891

Qy        277 GGGGGCCACCATCCGACAACATCACAAAACAGACCCCAGTCCAAGATAGACGTGCATAGGAA 336
              |||||||||||||||
Db        892 AGGTGCCACCATTGGAACATCACCAAACAGACCCAGTCTAAAATCGATGTCCACCGTAA 951

Qy        337 GGAGAACCGAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGTGCTC 396
              |||||||||| ||| ||||
Db        952 AGAAGGATGCGGGGGCTGCTGAGAAGTCATTGGAGAAGTTGCATGAGGGCAAAACGGTGAATGAACTTC 1011

Qy        397 CTCCGGCTTGTGAAGATGATCTTGGAGATTATGCATAAGAGAGGGCTAAGGACACCAAAACGGC 456
              |||| |||||
Db       1012 TGCGCGCTTGTGAAGTCTATTCTGGAGATTATGCGAGAAGAAGCTCAAGTATAAAAATTCAC 1071

Qy        457 TGACGAGGTTCCCCTGAAGATCCTGGCCCCATAATAACTTGTAGGGCGTCTCATTGGCAA 516
              || ||
Db       1072 AGAAGAGATCCCCCTGAGGAGATTTTAGCTCATAATAACTTGTTGGGACGTCTTATTGGTAA 1131

Qy        517 GGAAGGACGGAACCTGAGGAAGGTAGAGCAAGAGATACCGGAGAAAAATCACCATCTCCTC 576
              ||||
Db       1132 AGAAGGAAGAAGCAATCTTAAAAAAATTGAGCAAGAACGACACAGACACTAAAAATCACGATATCTCC 1191

Qy        577 GTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA 636
              ||||||||||||
Db       1192 ATTCAGGAATTGACGCGTCGTATAATCCAGAAACGCAGCTATTACAGTTAAAGGCAATGTTGA 1251

Qy        637 GAATTGTTGCAGGCGACGGGCCGACGGCAGAAATAATGAGAGAAAGAAGTTCGGGAGGCCTATGAGAATGA 696
```

```
Db       1252 GACATGTGCCAAAGCTGAGGAGGAGATCATGAGAAAATACAGGGAGTCTTATGAAAATGA 1311
              ||||||  ||||||||||  ||||||||||||||||||||||||||||||||| ||||||
Qy        697 TGTGGGCTGCCATGA-----GCTCTCACCTGACGCCCTGAACCTGGCTGCTGTAGG 750

Db       1312 TATTGCTTCTATGAATTCTCAAGCACATTAATTCCTGGATTAAATCTGAACGGCCTTGGG 1371
              ||| | ||||||  ||
Qy        751 TCTTTTCCCAGTTCATCCAGCATCAGTCCCGCCGCTCCGCGCTCCCACCTCAGCAGCCTGC 810

Db       1372 TCTGTTCCCACCCAGTTGCACGGATGCCACTTCCAGGCGGGCCCCTCCAGGGCCCATGAC 1431
              ||| |||||||||||
Qy        811 TCCCTATAGCTCCTTTATGCAGGGGTGCAGGCTCCGAGCGAGGAGATGGTGCAGTGTTTATCCCGC 870

Db       1432 TCCTCCTACCCGCAGTTGAGCAATCAGAATCAGGAGACTGTTCATCAGTTTATCCCAGC 1491
              |||||||||||||| |
Qy        871 CCAGGCAGTGGGGCCCATCATCGGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCGTT 930

Db       1492 TCTATCAGTCGGTGCCATCATCGGGCCATCATCGGGCAGCAGCTTTCTCGCTT 1551
              |||||||||||||||||||||
Qy        931 TGCCAGGCGCCTCCATCAAGATGCACCACCCGAAACACCTGACTCCAAAGTTCGTATGGT 990

Db       1552 TGCTGGAGCTTCAATTAAGATTGCTCCAGCGGCCAGCACCCAGCAGGCAGTGCTAAAGTGAGGATGGT 1611
              |||||||| |||||||
Qy        991 TATCATCACTGACCACCGCCAGAGCCCAGGGGCCCAATTCAAGGCTCAGGAAGAATCTATGGCAAACT 1050

Db       1612 GATTATCACTGACCACCAGACAGAGGCTCAGTTCAAGGCTCAGGAAGAATTTATGATGGAAAAT 1671
              ||||||||||||||||| |||
Qy       1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAGAGCCCACATACGTGT 1110

Db       1672 TAAAGAGAAAACTTTGTTAGTCCTAAAGAGGTGAAAACTTGAGCTCATATCAGAGT 1731
              ||||||||| ||| | |||||||
Qy       1111 GCCAGCATCAGCAGCTGCCGGGTCATTGGCAAAGTGGAAAAACGGTGAACGAGTGCA 1170

Db       1732 GCCATCCTTTGCTGCTGCAGAGTTATTGGAAAAGAGGCAAAACGGTGAATGAACTTCA 1791
              |||| ||||||| ||
Qy       1171 GAATTTGACGCGAGCTGAGGTCGTGAGTACCAAGAGACCACCCTGATGAGAAGACCA 1230

Db       1792 GAATTTGTCAAGTGCAGAAGTGTGTGTCCCTCGTGAGACCTGTGACCAGAGAATGAGACCA 1851
              ||||||| ||||||||| ||||
Qy       1231 GGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGTGGCTCAACGGAAGATCCG 1290

Db       1852 AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTTCTATCGTTGCCAGGTTGCCCAGAGAAAAATTCA 1911
              |||| | |||||| |||||
Qy       1291 AGACATCCTGGCCCAGTTAAGCAGCATCAGCAGCATCAGAAG 1327

Db       1912 GGAAATTCTGACTCAGGTAAAGCAGCACCAACAACAG 1948
```

```
RESULT 14
US-09-606-421B-175
;  Sequence 175, Application US/09606421B
;  Patent No. 6531315
;  GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT: Fan, Liqun
;  APPLICANT: Kalos, Michael D.
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT: Hosken, Nancy
;  APPLICANT: Fanger, Gary R.
;  APPLICANT: Li, Samuel X.
;  APPLICANT: Wang, Aijun
;  APPLICANT: Skeiky, Yasir A.W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C9
;  CURRENT APPLICATION NUMBER: US/09/606,421B
;  CURRENT FILING DATE: 2000-06-28
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
;  SEQ ID NO 175
;  LENGTH: 4181
;  TYPE: DNA
;  ORGANISM: Homo sapiens
```

```
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3347)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3502)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3506)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3520)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3538)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3549)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3646)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3940)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3968)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (3974)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (4036)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (4056)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (4062)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (4080)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (4088)
;    OTHER INFORMATION: n=A,T,C or G
;    FEATURE:
;    NAME/KEY: unsure
;    LOCATION: (4115)
;    OTHER INFORMATION: n=A,T,C or G
US-09-606-421B-175

Query Match         34.0%;   Score 580.6;  DB 4;   Length 4181;
Best Local Similarity   66.1%;
Matches  857;  Conservative    0;  Mismatches  434;  Indels    6;  Gaps    1;
```

```
Qy     97 CCCCGATGAGCAGATAGCACAGGGACCTGAGAATGGCGGCCGAGGGGGCTTTGGCTCTCG  156
             ||  ||   ||| ||| | ||||  || | ||   |   ||||||  | | || ||| |
Db    712 CCCTGATGAAATGGCCGCCAGCAAACCCCTTGCGAGCAGCCCCGAGGTGCGCCGGGGGCT  771

Qy    157 GGGTCAGCCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCTCCTCAAGGCAGCAGCCAAGT  216
             |||| ||    | |||| ||   ||  | |||        || ||||||||||||
Db    772 TGGGCAGAGGGGCTCCTCCAAGGCAGGGGTCTCCAAGGATCCGTATCCAAGCAGCAGAAACCATG  831

Qy    217 GGACATCCCCTCGGCTCCTGGTGCCCACCCAGTATGTGGTGCCATTATTGGCAAGGA  276
             ||||| ||   |||| |  | |||||| || | ||||  || | ||   ||||||||
Db    832 TGATTTGCCTCTGCGCCTGGTTCCCACCCAATTTGTTGGAGCCATCATAGAGGAAAAGA  891

Qy    277 GGGGGCACCATCCGCAACATCACAAAACAGACCCAGTCAAGATAGACCTGCATAGGAA  336
             | |||   || |||| | | |  | ||||  || ||   |||||| |||  |||||||
Db    892 AGGTGCCACCATTCGGAACATCACCAAACAGACCCAGTCTAAAATCGATGTCCACCGTAA  951

Qy    337 GGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTC  396
             ||| || | ||||| | ||||  ||  || |||   | |||   ||||  ||||||| ||
Db    952 AGAAAATGCGGGGGTCGTGAGACTGCGATTACTATCCTCTCTACTCCTGAAGGCACCTC  1011

Qy    397 CTCCGCTTGTAAGAATCCTGGCCTGAAGATCTTGTAGAGCGTCTCATTGGCAA  456
             || || ||| | || ||||||| |||||||  | | | |||| |  || |||||
Db   1012 TGCGGCTTGTAAGTCTTCATTCTGGAGATTATGCAAGAAGCTCAAGATATAAAATTCAC  1071

Qy    457 TGACGAGGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCATTGGCAA  516
             ||   ||||| || ||||   | |||||||||| | || | | ||  ||| |||||||
Db   1072 AGAAGAGATCCCCTTGAAGATTTAGCTCATCATAATAACTTTGTTGGACGTCTTATTGGTAA  1131

Qy    517 GGAAGGACGAACTGAAGAAGTAGAGCAAGATACCGAAGACAAAAATCACCATCTCCTC  576
             || |||||| | ||| || ||||   |||  |  || ||| |   | |  ||| |||
Db   1132 AGAAGGAAGAAATCTTAAAAAAATTGAGCAAGACACTCAGAGACACTAAAATCACGATATCTCC  1191

Qy    577 GTTGCAAGACCTTACCCTTTACAACCCTGAGCAGGGGGCCATCGA  636
             | |||||  | | |||  ||  | |||| ||   |||||||| ||
Db   1192 ATTGCAGGATTGACGCTGTATAATCCAGAACGACTATTACAGTTAAAGGCAATGTTGA  1251

Qy    637 GAATTGTTGCAGGGCCGAGCAGGAGAAATAATGAAGTTCGGGAGGCCTATGAGAATGA  696
             ||||| ||||| || |||| | || |  | |  |||  |||||| |  |||||| |||
Db   1252 GACATGTGCCAAAGTGAGGAGGAGATCATGAAGCATGAAGAAAATCAGGAGTCTTATGAAAATGA  1311

Qy    697 TGTGCTGCCATGA------GCTCTCACCTGATCCCTGGCCTCGTAGG  750
             | | ||||||||       |||||||  |||  | ||| ||  |||
Db   1312 TATTGCTTCTATGAATCTTCAAGCACATTTAATTCCTGGATTAAATCTGAACGCCCTTGGG  1371

Qy    751 TCTTTTCCCAGCTTCATCAGCGAGTCCCGCCGCAGCGTACTGGGGCTGC  810
             || || |||| |||| ||||| || ||| | | | ||||  ||||| ||
Db   1372 TCTGTTCCCACCCACTTCAGGGATGCCAACTTCCCACCTCCAGCAGGATGTGCAGGTGTTTATCCCCGC  1431

Qy    811 TCCCTATAGCTCCTTTATGCAGGGTCCCCGC  870
             || |  | | || || ||||| | || |||
Db   1432 TCCTCCCTACCCGAGTTTGAGCAATCAGAAACGGAGACTGTTCATCAGTTTATCCCAGC  1491

Qy    871 CCAGGCAGTGGGCGCCATCATCGGCAAGAACAGCTCTCCCGGTT  930
             |||||| || || || ||||||||| ||||  || |  ||||||
Db   1492 TCTATCAGTCGGTGCCATCATCGGCACCATCAAGCAGCTTTCTCGCTT  1551

Qy    931 TGCCAGCGCCTTCCATCAAGATTGCACACCGGAAGTTCTATGGT  990
             ||||||||| || || || || ||| | || || || |||||||
Db   1552 TGCTGGAGCTTCAATTAAGATTGCTCCAGCACCAGCAGATGCTAAAGTGCTAAAGTGCTAAAGTGGT  1611

Qy    991 TATCATCACTGGACCGCCAGAGCCCAATTCAAGGCTCAGGAGAATCTATGCAAACT  1050
             || |  ||| |||| | ||| || |||   ||||||||| ||| |   |||| |
Db   1612 GATTATCACTGGACCACCAGAGCTCAGGTCAGGAGAAGAATTTATGGAAAAT  1671

Qy   1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGGGAAGCTGAGAGCCCACATACGTGT  1110
             |||  |||  |  | ||| |||   | |||  | | ||| || | || |  ||   | |
Db   1672 TAAAGAAGAAAACTTTGTTAGTCCTAAAGAAGAGGTGAAACTTGAAGCTCATATCAGAGT  1731

Qy   1111 GCCAGCCATCAGCAGCTGGCCGGTCATTGGCAAAAACGGTGAACGAGTTGCA  1170
             ||||  ||| |  | || |||   | ||||| | | |  || || |||| ||
Db   1732 GCCATCCTTTGTCTGTGGCAGAGTTATTGGAAAGAGGAGAAACGGTGAATGAACTTCA  1791
```

```
Qy     37 AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTCGAAGGCTCCTACAT  96
             | ||||  | | ||| ||| | | | |||   | ||   ||  | |||||||| | ||||
Db    652 AGCCACTAGACAAACTGAATGGAATTCAGTTAGAGAATTCACCTTGAAAGTAGCCTATAT  711
```

```
Qy   1171  GAATTTGACGGCCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCTGATGAGAACGACCA  1230
           ||||||| || ||| |  ||   |    ||  |||  || |||   ||||| ||  ||
Db   1792  GAATTTGTCAAGTGCAGAGTTGTTGTCCCTCGTGACCAGCACCTGATCAGATGAGAATGACCA  1851

Qy   1231  GGTCATCGTGAAAATCATCGGACATTTTATGCGAGTCAGTCAGATGGCTCAACGGAAGATCCG  1290
           ||  ||| || ||  ||   ||  || |  || |  ||||  ||      ||  ||  |  ||
Db   1852  AGTGGTTGTCAAATAATAACTGGTCACTTCTATGTTGCCAGGTTGCCCAGAGAAAATTCA  1911

Qy   1291  AGACATCCTGGCCCAGTTAAGCAGCAGCATCAGAAG  1327
           || ||| | ||   ||| ||  |||||  |  |||
Db   1912  GGAAATTCTGACTCAGGTAAAGCAGCAGCACCAACAACAG  1948

RESULT 15
US-09-061-709-8
; Sequence 8, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-8

Query Match         29.5%;  Score 504;  DB 3;  Length 3283;
Best Local Similarity 63.6%;  Pred. No. 8.4e-134;
Matches  849;  Conservative   0;  Mismatches  425;  Indels  60;  Gaps  3;

Qy     37  AGCCATCATGAAGCTGAGTGATGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCCTACAT   96

Db    474  AGCCATGGAGGCTAAGCGCGGGACATCAGTCAGTTTGAGAACTACTCCTTCAAGATTTCCTACAT  533

Qy     97  CCCGATGAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGCCCCCCTTCGCCCCAGCCGTGG   156

Db    534  CCCGGATGA------AGAGGTGAGCTCCCCTGTCCAGCGGCGGGGCCCCAGCCAGCAGCAGT   584

Qy    157  GGGTCAGCCCGCCCGCCAGGGACTCACCTGTGGCAGCGGCGGGGCCCCAGCCAAGCAGCAGT   216

Db    585  GGACCACTCTTCCCGGGAGCAAGGCCACCGCACTTCTCAGGCCAGACAGAT            644

Qy    217  GGACATCCCCCTTCGGCTCCGGATCCTGGTCCCACCCAGTATGTGGTGCCATTATTGGCAAGGA   276

Db    705  TGTACAACAGGTGAAGCAGTGAAAGAGCAGCAGAAATACCCTCAGGGAGTCGCCTCACAGCGCAG  764

Qy    277  GGGGGCCACCATCCGCAACATCACAAAACAACAGTCCAAGATAGACGTGCATAGGAA   336

Db    705  GGGCTTGACCATAAAGAGACACTCACTGACGACACCCAGTCCCGGGTAGATAATCCATAGAAA   764

Qy    337  GGAGAACGCAGGTGCAGTGAAAAAGCCATCAGTGTGCACTCCACCCCCTGAGGGCTGCTC    396

Db    765  AGAGAACTCTGAGAGCTGCAGAGAAGCCATGTCACCATCCATGCAGAGGGGGACTTC    824

Qy    397  CTCCGCTTGTGAAGATGTCTTGGAGATTATGCATAAAGAGGTAAGGACACCAAAACGGC    456

Db    825  TGAAGCATCGCCGCATGATTCTTGAAATCATGCGAGCAAAGAGGCCAGATGAGAACCAAACTAGC  884

Qy    457  TGACGAGGTTCCCCTGCGAGAGATCCTGGCCCCCATAATAACTTTGCTGCCATTGGCAA   516

Db    885  CGAAGAGATTCCTCGAAAATTCTGGCACACAATGGCTTGGTTGGAAGACTGATTGGAAA   944

Qy   517  GGAAGGACGGAACCTGAGAGAGGTGAGCTAGAGCCAAGATACCGAGACAAAATCACCATCCTC   576

Db   945  AGAAGGCCAGAGAAATTGAAGAAGAAATTGAACATGAAACAGGGACCAAGATAACAATCTTCATC 1004

Qy   577  GTTGCAAGACCTTACCCTTTACAACCTGAGAGGACCATCACTGTGAAGGGGGCCATCGA   636

Db  1005  TTTGCAGGATTTGAGCATATACAACCCGGAAAGAACCATCACGTGAAGGGCACCAGTTGA  1064

Qy   637  GAATTGTTGCAGGGCCGAGCAGGAGAAATAATGAAGAAGTTGGGGAGCCTATGAGAGAATGA   696

Db  1065  GGCCTGTGCCAGTGCTGAGATAGAGATTATGAAGAGTCGCGTGAGCCTTTGAAAATGA  1124

Qy   697  TGTGGCTGCCATGAGCTCTCACCTGATCCCTGGCCTGAACCTGGCTCGCTGTAGGTCTTT   756

Db  1125  TATGCTGCGTGTTAACACCCA---------------CTCCGGATACTTCTCC  1161

Qy   757  CCCAGTTCATCCAGCGCAGTCCCGCCTTCCCAGCAGCGTTACTGGGGCTGCTCCCTA   816

Db  1162  AGCCTGTACCCCCATCACCAGTTGCCCGTTCCCGCATCATCACT------------  1207

Qy   817  TAGCTCCTTTATGCAGGCTCCCGAGCAGGAGGTGGGTGCAGGTGTTTATCCCGCCCAGGC   876

Db  1208  ----------CTTATCCAGAGCAGGAGGATTGTTGAATCTTCCATCCCAACCCAGGC  1253

Qy   877  AGTGGGCGCCATCATCGGCCAAGAGAAGGGCAGCACATCAAACAGCTCTCCCGGTTGCCAG   936

Db  1254  TGTGGGGCCATCATCGGGAGGAAGGGGCACACCATCAAACAGCTGCGAGATTCGGCCGG  1313

Qy   937  CGCCTCCATCAAGATTGCACCACCCGAAAACACCTGACTCAAAGTTCGTATGGTTATCAT   996

Db  1314  AGCCCTCTATCAAGATTGCCCCCTGCGGAAGGCCCAGACGTCAGCGAAAGGATGGTCATCAT  1373

Qy   997  CACTGGACGCCAGAGCCCATTCAAGGCTCAGGGAAGAATCATCGCAAACTCAAGGA  1056

Db  1374  CACCGGGCCACCGGAAGCCCAGTTCAAGGCCCAGTTCAAGGCGACGACGGATCTTTGGGAAACTGAAAGA  1433

Qy  1057  GGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCCACATACGTGTGCCAGC  1116

Db  1434  GGAAAAACTTCTTTAACCCCAAAGAACGAAGTGAAGCTCGAAGCCATATCAGAGTGCCCTC  1493

Qy  1117  ATCAGCAGCTGCTGCCGGTCATTGGCCACCAGTTGGAAAAACGTGAACGAGTGCAGAATTT  1176

Db  1494  TTCCACAGCTGCGCGGTGATTGGCCTGACAAGGTTGGCAAGACCGTGACGAACTGCAGAACTT  1553

Qy  1177  GACGGGCAGCTGAGGTGGTAGTACCAAGAGACACCCTGATGAGAACGACCAGGTCAT  1236

Db  1554  AACCAGTGCAGAAGTCATCGTGCCTCGTGACCAGCCCAAACGGCCAGTGAAGAAATGAGGAAGTGAT  1613

Qy  1237  CGTGAAAATCATCGGACATTTCTATGCCAGTCAGTAGGCTCAACGGAAGATCCGAGACAT  1296

Db  1614  CGTCAGAATTATCGGGCACTTCTTTGCTGTAGCCAGATCTCTAGCAGCGCCAAGATCAGGGAAAT  1673

Qy  1297  CCTGGCCCAGGTTAAGCAGTGAAGCTGAGCATCAGAAGGACAGAGAGTAACCAGGCACACGGAG  1356

Db  1674  TGTACAACAGGTGAAGCAGTGAAAGAGCAGCAGAAATACCCTCAGGGAGTCGCCTCACAGCGGCAG  1733

Qy  1357  GAAGTGACCAGCCC  1370

Db  1734  CAAGTGAGGCTCCC  1747
```

Search completed: July 13, 2004, 21:42:51
Job time : 141 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: July 13, 2004, 14:54:07 ; Search time 6722 Seconds
                                          (without alignments)
                                          11013.077 Million cell updates/sec

Title: US-09-270-437D-5
Perfect score: 1708
Sequence: 1 aggacgcgcgcgaccgc........atttcctcaggtttaaaa 1708

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151699 5 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
          1:   gb_ba:*
          2:   gb_htg:*
          3:   gb_in:*
          4:   gb_om:*
          5:   gb_ov:*
          6:   gb_pat:*
          7:   gb_ph:*
          8:   gb_pl:*
          9:   gb_pr:*
         10:   gb_ro:*
         11:   gb_sts:*
         12:   gb_sy:*
         13:   gb_un:*
         14:   gb_vi:*
         15:   em_ba:*
         16:   em_fun:*
         17:   em_hum:*
         18:   em_in:*
         19:   em_mu:*
         20:   em_om:*
         21:   em_or:*
         22:   em_ov:*
         23:   em_pat:*
         24:   em_ph:*
         25:   em_pl:*
         26:   em_ro:*
         27:   em_sts:*
         28:   em_un:*
         29:   em_vi:*
         30:   em_htg_hum:*
         31:   em_htg_inv:*
         32:   em_htg_other:*
         33:   em_htg_mus:*
         34:   em_htg_pln:*
         35:   em_htg_rod:*
         36:   em_htg_mam:*
         37:   em_htg_vrt:*
         38:   em_sy:*
         39:   em_htgo_hum:*
         40:   em_htgo_mus:*
         41:   em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1703	99.7	1708	6	AR171864	AR171864 Sequence
2	1703	99.7	1708	6	AR343073	AR343073 Sequence
3	1703	99.7	1708	6	BD209924	BD209924 Isolated
4	1669.2	97.7	1946	6	AR171866	AR171866 Sequence
5	1669.2	97.7	1946	6	AR343075	AR343075 Sequence
6	1669.2	97.7	1946	6	BD209926	BD209926 Isolated
7	1653.2	96.8	2130	9	AF117106	AF117106 Homo sapi
8	1631.8	95.5	2381	9	AF198254	AF198254 Homo sapi
9	1389	81.3	2780	6	BD127811	BD127811 Primer fo
10	1389	81.3	2780	9	AK074915	AK074915 Homo sapi
11	1165	68.2	2224	6	AR160244	AR160244 Sequence
12	1162.6	68.1	2223	10	AF061569	AF061569 Mus muscu
13	1161.6	68.0	2444	10	BC051679	BC051679 Mus muscu
14	1112.4	65.1	2105	10	AF541940	AF541940 Rattus no
15	914.8	53.6	2021	5	AF026527	AF026527 Gallus ga
16	732.8	42.9	833	6	BD124998	BD124998 Primer fo
17	732.8	42.9	833	6	BD126684	BD126684 Primer fo
18	627	36.7	1740	10	AB046173	AB046173 Mus muscu
19	627	36.7	2140	10	BC045138	BC045138 Mus muscu
20	627	36.7	4030	10	BC049082	BC049082 Mus muscu
21	588	34.4	2011	5	AF161270	AF161270 Danio rer
22	588	34.4	2612	5	BC045873	BC045873 Danio rer
23	582.2	34.1	1740	5	BD275902	BD275902 COMPOUNDS
24	582.2	34.1	1740	6	AR220687	AR220687 Sequence
25	582.2	34.1	1740	6	AR281251	AR281251 Sequence
26	582.2	34.1	1740	6	AX365954	AX365954 Sequence
27	582.2	34.1	1743	6	AX366054	AX366054 Sequence
28	582.2	34.1	1764	6	AX366057	AX366057 Sequence
29	582.2	34.1	4171	6	AX366035	AX366035 Sequence
30	582.2	34.1	4171	9	AF117108	AF117108 Homo sapi
31	580.6	34.0	4155	9	HSU76705	U76705 Human putat
32	580.6	34.0	4159	6	AR171863	AR171863 Sequence
33	580.6	34.0	4159	6	AR343072	AR343072 Sequence
34	580.6	34.0	4159	6	BD209923	BD209923 Isolated
35	580.6	34.0	4181	6	BD275767	BD275767 COMPOUNDS
36	580.6	34.0	4181	6	AR220552	AR220552 Sequence
37	580.6	34.0	4181	6	AR255546	AR255546 Sequence
38	580.6	34.0	4181	6	AR281116	AR281116 Sequence
39	580.6	34.0	4181	6	AX333233	AX333233 Sequence
40	580.6	34.0	4181	6	AX365782	AX365782 Sequence
41	580.6	34.0	4181	9	HSU97188	U97188 Homo sapien
42	580.6	34.0	4602	9	AX397963	AX397963 Sequence
43	571.8	33.5	2161	5	AF055923	AF055923 Xenopus 1
44	571.8	33.5	2405	5	AF064634	AF064634 Xenopus 1
45	562.6	32.9	2031	5	BC057700	BC057700 Xenopus 1

ALIGNMENTS

RESULT 1
AR171864
LOCUS   AR171864   1708 bp   DNA   linear   PAT 17-DEC-2001
DEFINITION Sequence 5 from patent US 6297364.
ACCESSION   AR171864
VERSION   AR171864.1 GI:17910814
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1708)
  AUTHORS   Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
            and Old,L.J
  TITLE   Isolated nucleic acid molecule encoding cancer associated antigen,
            the antigen itself, and uses thereof

JOURNAL   Patent: US 6297364-A 5 02-OCT-2001;
FEATURES   Location/Qualifiers
     source   1..1708
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match   99.7%; Score 1703; DB 6; Length 1708;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1708; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1  AGGGACGCTGCCGCACCGCCCAGTTTACCCCGGGGAGCCATCATGAAGCTGATGGCCA   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  AGGGACGCTGCCGCACCGCCCAGTTTACCCCGGGGAGCCATCATGAAGCTGATGGCCA   60

Qy     61  CCAGTTGAGAACCATGCCTCTGAAGGTCTCTACATCCCGATGACGAGATAGCACAGGG  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CCAGTTGAGAACCATGCCTCTGAAGGTCTCTACATCCCGATGACGAGATAGCACAGGG  120

Qy    121  ACCTGAGAATGGGCGCCGAGGGGCTTTGGCTTCGGGTCAGCCCCGCCCAGGGCTCACC  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  ACCTGAGAATGGGCGCCGAGGGGCTTTGGCTTCGGGTCAGCCCCGCCCAGGGCTCACC  180

Qy    181  TGTGGCAGCGGGGGGCCCCAGCCAAGCAGCAGCAAGTGACATCCCCTTCGGCTCCTGGT  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TGTGGCAGCGGGGGGCCCCAGCCAAGCAGCAGCAAGTGACATCCCCTTCGGCTCCTGGT  240

Qy    241  GCCACCCAGTATGTGGGTGCCATTATTGGCAAGAGCACGGGGGCCACATCCGCAACATCAC  300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GCCACCCAGTATGTGGGTGCCATTATTGGCAAGAGCACGGGGGCCACATCCGCAACATCAC  300

Qy    301  AAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAAGCAGTGCAGTGAAAA  360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  AAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAAGCAGTGCAGTGAAAA  360

Qy    361  AGCCATCAGTGTGCACTCACCCCTGAGGGCTGCTCCTCCGGCTTGTCTGTGAGGGATCTTGGA  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AGCCATCAGTGTGCACTCACCCCTGAGGGCTGCTCCTCCGGCTTGTCTGTGAGGGATCTTGGA  420

Qy    421  GATTATGCATAAAGAGAGCTAAGGACACCAAAAACGGCTGACGAGGTTCCCCTGAAGATCCT  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GATTATGCATAAAGAGAGCTAAGGACACCAAAAACGGCTGACGAGGTTCCCCTGAAGATCCT  480

Qy    481  GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGCGGAACCTGAAGAGGT  540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGCGGAACCTGAAGAGAGT  540

Qy    541  AGAGCAAGATACCGAGACAAAAATCACCATCTCCGTGTTGCAAGACCTTACCCTTTACAA  600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  AGAGCAAGATACCGAGACAAAAATCACCATCTCCGTGTTGCAAGACCTTACCCTTTACAA  600

Qy    601  CCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGGCCGAGCCAGGA  660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGGCCGAGCCAGGA  660

Qy    661  AATAATGAAGAAAGTTCGGAGAGCCCATGAGAATGATGTGGCTGCCATGAGCTCTCACCT  720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AATAATGAAGAAAGTTCGGAGAGCCCATGAGAATGATGTGGCTGCCATGAGCTCTCACCT  720

Qy    721  GATCCCCTGCCTGGCTGTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCCAGCGCAGTCCC  780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GATCCCCTGCCTGGCTGTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCCAGCGCAGTCCC  780

Qy    781  GCCGCCCTCCCAGCCAGCGGTTACTGGGCGTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGA  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GCCGCCCTCCCAGCCAGCGGTTACTGGGCGTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGA  840

Qy    841  GCAGGAGATGGTGCAGGTGTTTATCCCGCCAGTAGCCATCATCGGCAGCTCCGAAGAA  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  GCAGGAGATGGTGCAGGTGTTTATCCCGCCAGTAGCCATCATCGGCGGCCATCATCCGGCGGAAGAA  900

Qy    901  GGGGCAGCACATCAAACAGCTCTCCCGGTTTGCCAGGCCTCCATCAAGATTGCACCACC  960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  GGGGCAGCACATCAAACAGCTCTCCCGGTTTGCCAGGCCTCCATCAAGATTGCACCACC  960
```

```
Db    901  GGGGCAGCACATCAAACAGCTCTCCCGGTTTGCCAGGCCTCCATCAAGATTGCACCACC   960
Qy    961  CGAAACACCTGACTCCAAAGTTCGTATGGTATCATCACTGAGGTTATCATCACTGGACGCGCCAGAGCGCCAATT  1020
Db    961  CGAAACACCTGACTCCAAAGTTCGTATGGTATCATCACTGAGGTTATCATCACTGGACGCGCCAGAGCGCCAATT  1020

Qy   1021  CAAGGCTCAGGGAAGGAAGAATCATGCCAAACTCAAGGAGGAGAACTTCTTTGGTCCTCCAAGGA  1080
Db   1021  CAAGGCTCAGGGAAGGAAGAATCATGCCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGA  1080

Qy   1081  GGAAGTGAAGCTGGAGACCCCACATACGTGCCAGCATCAGCAGCTGGCCGCGGGTCATTGG  1140
Db   1081  GGAAGTGAAGCTGGAGACCCCACATACGTGCCAGCATCAGCAGCTGGCCGCGGGTCATTGG  1140

Qy   1141  CAAAGGTGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTGGTAGTACC  1200
Db   1141  CAAAGGTGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTGGTAGTACC  1200

Qy   1201  AAGAGACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAAATCATCCGGACATTTCTA  1260
Db   1201  AAGAGACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAAATCATCGGACATTTCTA  1260

Qy   1261  TGCCAGTCAGATGGCTCAACGAGCAAGTGACATCCGGAGAGATCCGGAGAAGCAGCACGA  1320
Db   1261  TGCCAGTCAGATGGCTCAACGAGCAAGTGACATCCGGAGAGATCCGGAGAAGCAGCACGA  1320

Qy   1321  TCAGAAGGGACAGAGTAACCAGGCCAGGAGTGACGAGAGTGACCAGCCCTCCTGTCC  1380
Db   1321  TCAGAAGGGACAGAGTAACCAGGCCAGGAGTGACGAGAGTGACCAGCCCTCCTGTCC  1380

Qy   1381  CTTNGAGTCAGGACAACAACGGGACGAGAAATCGAGAGTGTGCTCTCCCGGCAGGCCTGA  1440
Db   1381  CTTNGAGTCAGGACAACAACGGGACGAGAAATCGAGAGTGTGCTCTCCCGGCAGGCCTGA  1440

Qy   1441  GAATGAGTGGGAATCGGGACACNTGGGCGCGGGTTGCCCACTTGATT  1500
Db   1441  GAATGAGTGGGAATCGGGACACNTGGGCGCGGGTTGCCCACTTGATT  1500

Qy   1501  GAGAAAGATGTTCAGTGAGGAGAACCCTGATCTNTCAGCCCAAACACCCAACACCCAATTGC  1560
Db   1501  GAGAAAGATGTTCAGTGAGGAGAACCCTGATCTNTCAGCCCAAACACCCAACACCCAATTGC  1560

Qy   1561  CCAACACTGTNTGCCCTCGGGGTGTCAGAAATTNTAGGCCAAGGCACTTTTAAACGTGG  1620
Db   1561  CCAACACTGTNTGCCCTCGGGGTGTCAGAAATTNTAGGCCAAGGCACTTTTAAACGTG  1620

Qy   1621  ATTGTTTAAAGAAGTCTCCAGGCCCCAAGAGGGTGATCACACCTCAGTGGGAAGA  1680
Db   1621  ATTGTTTAAAGAAGTCTCCAGGCCCCAAGAGGGTGATCACACCTCAGTGGGAAGA  1680

Qy   1681  AAAATAAAATTTCCTTCAGTTTTAAAA  1708
Db   1681  AAAATAAAATTTCCTTCAGTTTTAAAA  1708
```

RESULT 2
AR343073
LOCUS   AR343073   1708 bp   DNA   linear   PAT 17-AUG-2003
DEFINITION Sequence 5 from patent US 6576756.
ACCESSION   AR343073
VERSION   AR343073.1 GI:33738475
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1708)
  AUTHORS   Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
            and Old,L.J.
  TITLE   Isolated nucleic acid molecule encoding cancer associated antigen,
            the antigen itself, and uses thereof
  JOURNAL   Patent: US 6576756-A 5 10-JUN-2003;
FEATURES   Location/Qualifiers
     source   1..1708

```
                    /organism="unknown"
                    /mol_type="genomic DNA"
ORIGIN

Query Match      99.7%;   Score 1703;  DB 6;  Length 1708;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1708;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy      1  AGGGACGCGCCGCACCGCCCAGTTTACCCCGGGAGCTCCTACATGAAGCTGAATGCGCCA   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  AGGGACGCGCCGCACCGCCCAGTTTACCCCGGGAGCTCCTACATGAAGCTGAATGCGCCA   60

Qy     61  CCAGTTGGAGAACCATGCCCTGAAGGTCTCTACATCCCGATGAGCAGATAGCACAGGG   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CCAGTTGGAGAACCATGCCCTGAAGGTCTCTACATCCCGATGAGCAGATAGCACAGGG   120

Qy    121  ACCTGAGAATGGGCGCCAGGGCCCCAGCTTTGGCTTCTCGGGGTCAGCCCCGCCAGGGCTCACC  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  ACCTGAGAATGGGCGCCAGGGCCCCAGCTTTGGCTTCTCGGGGTCAGCCCCGCCAGGGCTCACC  180

Qy    181  TGTGGCAGCGGGGCCCCAGCCAAGCAGCAGCAAGCAGCATCCCCCTTCGGCTCCTCCTGGT  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TGTGGCAGCGGGGCCCCAGCCAAGCAGCAGCAAGCAGCATCCCCCTTCGGCTCCTCCTGGT  240

Qy    241  GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCAC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCAC  300

Qy    301  AAAACAGACCCAGTCAAGATAGAGCTGCATAGGAAGGAGAACGCAGGTGCAGGTGCGAAAA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  AAAACAGACCCAGTCAAGATAGAGCTGCATAGGAAGGAGAACGCAGGTGCAGGTGCGAAAA  360

Qy    361  AGCCATCAGTGCACTCCACCCTGAGGGCTGCTCTCCGCTTGTTGTAAGATGATCTTGGA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  AGCCATCAGTGCACTCCACCCTGAGGGCTGCTCTCCGCTTGTTGTAAGATGATCTTGGA  420

Qy    421  GATTATGCATAAAGAGGCTAAGGACACCAAAAACGGCTGACGAGGTTCCCCTGACGAGATCCT  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GATTATGCATAAAGAGGCTAAGGACACCAAAAACGGCTGACGAGGTTCCCCTGACGAGATCCT  480

Qy    481  GGCCCATAATAACTTGTAGGGCGTCCATTGGCAAGGAGGACGGAACCTGAAGGAAGGT  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  GGCCCATAATAACTTTGTAGGGCGTCCATTGGCAAGGAGGACGGAACCTGAAGGAAGGT  540

Qy    541  AGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTGCAAGACCTTACCCTTTACAA  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  AGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTGCAAGACCTTACCCTTTACAA  600

Qy    601  CCCTGAGGAGAACTCACTGTGAAGGGGCCATGAAGGGGCCATGAGATGATGTGGCTGCCATGAGCTCTCCACCT  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CCCTGAGGAGAACTCACTGTGAAGGGGCCATGAAGGGGCCATGAGATGATGTGGCTGCCATGAGCTCTCCACCT  660

Qy    661  AATAATGAAGAAGAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGAATGATGTGGCTGCCATGACTCCTCACT  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  AATAATGAAGAAGAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGAATGATGTGGCTGCCATGACTCCTCACT  720

Qy    721  GATCCCTGGCCTGAACCTGGCTGCTGCTGTAGGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GATCCCTGGCCTGAACCTGGCTGCTGCTGTAGGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCC  780

Qy    781  GCCGGCCTCCCAGCAGCGCGTTACTGGGGCTGCTCCCCGCTGCTCCCTATAGTCTCCTTTATGCAGGCTCCCGA  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  GCCGGCCTCCCAGCAGCGCGTTACTGGGGCTGCTCCCCGCTGCTCCCTATAGTCTCCTTTATGCAGGCTCCCGA  840

Qy    841  GCAGGAGATGGTGCAGGTGTTATCCCGCGCAGTGGGCGCCATCATCGGCCAAGAA  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  GCAGGAGATGGTGCAGGTGTTATCCCGCGCAGTGGGCGCCATCATCGGCCAAGAA  900

Qy    901  GGGGCAGCACATCAAACAGCTTCTCCCGGTTTGCCCAGCGGCCTCCATCAAGATTGCACCACC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  GGGGCAGCACATCAAACAGCTTCTCCCGGTTTGCCCAGCGGCCTCCATCAAGATTGCACCACC  960

Qy    961  CGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACGCCGCAGAGCCCCAATT  1020
```

```
Db    961  CGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACGCCGCAGAGCCCCAATT  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1021  CAAGGCTCAGGGAAGAGATCTATGGCAAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGA  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  CAAGGCTCAGGGAAGAGATCTATGGCAAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGA  1080

Qy   1081  GGAAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCAGCAGTCGGTCCGGGTCATTGG  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  GGAAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCAGCAGTCGGTCCGGGTCATTGG  1140

Qy   1141  CAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTGACGCGCAGCGAGTTGAGGTGGTAGTACC  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTGACGCGCAGCGAGTTGAGGTGGTAGTACC  1200

Qy   1201  AAGAGACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTA  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  AAGAGACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTA  1260

Qy   1261  TGCCAGTCAGATGGCTCAACGGAGAATCCGGAAGAATCCGGAAGACACCAGGTTAAGCAGCAGCA  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  TGCCAGTCAGATGGCTCAACGGAGAATCCGGAAGAATCCGGAAGACACCAGGTTAAGCAGCAGCA  1320

Qy   1321  TCAGAGGGACAGAGTAACCAGGGCCCCAGGACGCGGAGAAGTGACCAGCCCCTCCCGTCC  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  TCAGAGGGACAGAGTAACCAGGGCCCCAGGACGCGGAGAAGTGACCAGCCCCTCCCGTCC  1380

Qy   1381  CTTNGAGTCCAGGACAACACACGGGCAGAAATCGAGAATCGAGAGTGTGTCTCTCCCCGGCAGGCCTGA  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  CTTNGAGTCCAGGACAACACACGGGCAGAAATCGAGAATCGAGAGTGTGTCTCTCCCCGGCAGGCCTGA  1440

Qy   1441  GAATGAGTGGGAATCCGGGACACNTGGGCCGGGCTGTAGATCAGGTTGCCCACTTGATT  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  GAATGAGTGGGAATCCGGGACACNTGGGCCGGGCTGTAGATCAGGTTGCCCACTTGATT  1500

Qy   1501  GAGAAAGATGTTCCAGTGAGGAACCCTGATCNTNCAGCCCCAAACACCCCAATTGGC  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501  GAGAAAGATGTTCCAGTGAGGAACCCTGATCNTNCAGCCCCAAACACCCCAATTGGC  1560

Qy   1561  CCAACACTGTNTGCCCCTCGGGGTGTCAGAAATTNTAGCGGCAAGGCACTTTTAAACGTGG  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561  CCAACACTGTNTGCCCCTCGGGGTGTCAGAAATTNTAGCGGCAAGGCACTTTTAAACGTGG  1620

Qy   1621  ATTGTTTAAAGAAGTCTCCAGGCCCACCAAGAGGGTGGATCAAGAACTCAGTGGGAAGA  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621  ATTGTTTAAAGAAGTCTCCAGGCCCACCAAGAGGGTGGATCAAGAACTCAGTGGGAAGA  1680

Qy   1681  AAAATAAAATTTCCTTCAGGTTTTAAAA  1708
           ||||||||||||||||||||||||||||
Db   1681  AAAATAAAATTTCCTTCAGGTTTTAAAA  1708
```

```
RESULT 3
BD209924
LOCUS       BD209924    1708 bp    DNA    linear    PAT 17-JUL-2003
DEFINITION  Isolated nucleic acid molecules encoding cancer-associated
            antigens, these antigens and method of using the same.
ACCESSION   BD209924
VERSION     BD209924.1    GI:3019694
KEYWORDS    JP 2002512049-A/3.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1708)
  AUTHORS   Chen,Y.T., Gure,A., Tsang,S., Stockert,E., Jager,E., Knuth,A. and
            Old,L.J.
  TITLE     Isolated nucleic acid molecules encoding cancer-associated
            antigens, these antigens and method of using the same
  JOURNAL   Patent: JP 2002512049-A 3 23-APR-2002;
            LUDWIG INSTITUTE FOR CANCER RESEARCH
            OS    Homo sapiens (human)
            PN    JP 2002512049-A/3
COMMENT
```

```
PD   23-APR-2002
PF   16-MAR-1999 JP 2000545030
PR   17-APR-1998 US   09/061709
PI   YAO TSENG CHEN,ALI GURE,SOLAM TSANG,ELISABETH STOCKERT,ELKE
PI   JAGER,
PI   ALEXANDER KNUTH,LLOYD J OLD
PC   C12N15/09,A61K35/12,A61K39/00,A61K39/39,A61P35/00,C07K16/32,
PC   C12N1/15,
PC   C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/ PC
     574,
PC   C12N15/00,C12N5/00
CC   n=not determined(1384,1464,1533,1571,1595)
FH   Key             Location/Qualifiers
FT   source          1..1708
FT                   /organism='Homo sapiens (human)'.

FEATURES            Location/Qualifiers
     source         1..1708
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"

ORIGIN

Query Match          99.7%;   Score 1703;   DB 6;   Length 1708;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1708;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
```

```
Qy      1 AGGGACGCTGCCGCACCGCCCAGTTTACCCCGGGAGCCATCATGAAGCTGAATGGCCA    60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGGGACGCTGCCGCACCGCCCAGTTTACCCCGGGAGCCATCATGAAGCTGAATGGCCA    60

Qy     61 CCAGTTCGAGAACCATGCCCTGAAGGTCTCTACATCCCCGATGAGCAGATAGCACAGGG   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCAGTTCGAGAACCATGCCCTGAAGGTCTCTACATCCCCGATGAGCAGATAGCACAGGG   120

Qy    121 ACCTGGAATGGGCGCGGGGGCCCAGCCAGCAGCAAGCACGAGTGTTTGGCTTCGGGGTCACC   180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACCTGGAATGGGCGCGGGGGCCCAGCCAGCAGCAAGCAGTGTTTGGCTTCGGGGTCACC   180

Qy    181 TGTGGCAGCGGCGGGGGCCCCAGCCAGCAGCAGCAGCATCCCTTCGGCTTCCTGGT   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTGGCAGCGGCGGGGGCCCCAGCCAGCAGCAGCAGCATCCCTTCGGCTTCCTGGT   240

Qy    241 GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGAGGGGCCACCATCCGCAACATCAC   300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGAGGGGCCACCATCCGCAACATCAC   300

Qy    301 AAAACAGACCCAGTCAAGATAGACGTGCATAGGAAGGAAGAAACCAGGTGCAGTGAAAA   360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAAACAGACCCAGTCAAGATAGACGTGCATAGGAAGGAAGAAACCAGGTGCAGTGAAAA   360

Qy    361 AGCCATCAGTGTGCACTCCACCCTGAGGGCTGCTCTCCGCTTGTGAAGATGATCTTGGA   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGCCATCAGTGTGCACTCCACCCTGAGGGCTGCTCTCCGCTTGTGAAGATGATCTTGGA   420

Qy    421 GATTATGCATAAAGAGGCTAAGGACACCAAAAACGGCTGACGAGGTCCCCTGAAGATCCT   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GATTATGCATAAAGAGGCTAAGGACACCAAAAACGGCTGACGAGGTCCCCTGAAGATCCT   480

Qy    481 GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGCGGAAGCACGCGAAGAAGGT   540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGCGGAAGCACGCGAAGAAGGT   540

Qy    541 AGAGCAAGATACCGAGACAAAAATCACCATCCTCGTTGCAAGAACCTTACCCTTTACAA   600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGAGCAAGATACCGAGACAAAAATCACCATCCTCGTTGCAAGAACCTTACCCTTTACAA   600

Qy    601 CCCTGAGCAGGACCATCACTGTGAAGGGGCCATCGAGAAGGGCCATGAGAGATGATG   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCCTGAGCAGGACCATCACTGTGAAGGGGCCATCGAGAAGGGCCATGAGAGATGATG   720

Qy    661 AATAATGAAGAAAGTTCGGGAGGCCATGAGAGATGATGTGGCTGCCATGAGCTTCTCACCT   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATAATGAAGAAAGTTCGGGAGGCCATGAGAGATGATGTGGCTGCCATGAGCTTCTCACCT   720

Qy    721 GATCCCTGCCTGAACCTGGCTGCTGTCTTTCCCAGCTTCTTTCCCAGCTTCCAGCCAGTCCC   780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GATCCCTGCCTGAACCTGGCTGCTGTCTTTCCCAGCTTCTTTCCCAGCTTCCAGCCAGTCCC   780

Qy    781 GCCGGCCTCCCAGCAGCGTTACTGGGGCTTACTGCGGCTCCCCATAGTCTCTTTATGCAGGCTCCCGA   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCGCCTCCCAGCAGCGTTACTGGGGCTTACTGCGGCTCCCCATAGTCTCTTTATGCAGGCTCCCGA   840

Qy    841 GCAGGAGATGGTGCGAGGTGTTATCCCCAGGCAGTGGGCGCCATCATCGGCAAGAA   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GCAGGAGATGGTGCGAGGTGTTATCCCCAGGCAGTGGGCGCCATCATCGGCAAGAA   900

Qy    901 GGGGCAGCAGCATCAAAACAGCTCTCCCCGGTTTGCCAGGCGCCTCCATCAAGATTGCACCACC   960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GGGGCAGCAGCATCAAAACAGCTCTCCCCGGTTTGCCAGGCGCCTCCATCAAGATTGCACCACC   960

Qy    961 CGAAACACCTGACTTCCAAAGTTCGTATGGTTATCATCACTGACACGGCCCAGAGAGGCCCAATT   1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGAAACACCTGACTTCCAAAGTTCGTATGGTTATCATCACTGACACGGCCCAGAGAGGCCCAATT   1020

Qy   1021 CAAGGCTCAGGGAGAAGAATCATGCAAACTCAAGGAGGAGAACTTCTTTGGTCCAAGGA   1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAAGGCTCAGGGAGAAGAATCATGCAAACTCAAGGAGGAGAACTTCTTTGGTCCAAGGA   1080

Qy   1081 GGAAGTGAAGCTGGAAGACTCCGAGACCCACATACGTGTGCCAGCATCAGCAGCTCGGGTCATTGG   1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGAAGTGAAGCTGGAAGACTCCGAGACCCACATACGTGTGCCAGCATCAGCAGCTCGGGTCATTGG   1140

Qy   1141 CAAAGGTGAAAAACGGTTGAACCAGTTGCAGAATTTGACGGCAGCTGAGGTGTAGTACC   1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAAAGGTGAAAAACGGTTGAACCAGTTGCAGAATTTGACGGCAGCTGAGGTGTAGTACC   1200

Qy   1201 AAGAGACCAGACCCCTGATGAGACGACCAGTTCATCGTGAAAATCATCGGACATTTCTA   1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAGAGACCAGACCCCTGATGAGACGACCAGTTCATCGTGAAAATCATCGGACATTTCTA   1260

Qy   1261 TGCCAGTCAGATGGCTCAACGGAAGATCCGAGAACATCCTGGCCCAGTTAAGCAGCAGCA   1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGCCAGTCAGATGGCTCAACGGAAGATCCGAGAACATCCTGGCCCAGTTAAGCAGCAGCA   1320

Qy   1321 TCAGAGGGACAGAGTAACCAGCAGGCACCGGAAGTGACCAGCCCCTCCTGTCC   1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TCAGAGGGACAGAGTAACCAGCAGGCACCGGAAGTGACCAGCCCCTCCTGTCC   1380

Qy   1381 CTTNGAGTCCAGGACAACAACGGCAGAAATCGAGAGTGTGCTCTCCCCGGAGGCCTGA   1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CTTNGAGTCCAGGACAACAACGGCAGAAATCGAGAGTGTGCTCTCCCCGGCAGGCCTGA   1440

Qy   1441 GAATGAGTGGGAATCCGGGACACNTGGGCCGGGCTTGCCACACTTGATT   1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GAATGAGTGGGAATCCGGGACACNTGGGCCGGGCTTGCCACACTTGATT   1500

Qy   1501 GAGAAAGATGTTCAGTGAGGAACCCTGATCNTCAGCCCAATTGGC   1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAGAAAGATGTTCAGTGAGGAACCCTGATCNTCAGCCCAATTGGC   1560

Qy   1561 CCAACTGTNTGCCCCTCGGGGTGTCAGAAATNTAGCGCAAGGCACTTTAAACGTGG   1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCAACTGTNTGCCCCTCGGGGTGTCAGAAATNTAGCGCAAGGCACTTTAAACGTGG   1620

Qy   1621 ATTGTTTAAAGAAGCTTCCTTGCAAGACCTTACCCTTTACAA   1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ATTGTTTAAAGAAGCTTCCTTGCAAGACCTTACCCTTTACAA   1680

Qy   1681 AAAATAAAATTTCCTTCAGGTTTTAAAA   1708
          ||||||||||||||||||||||||||||
Db   1681 AAAATAAAATTTCCTTCAGGTTTTAAAA   1708
```

```
RESULT 4
AR171866
```

```
LOCUS       AR171866                1946 bp    DNA    linear    PAT 17-DEC-2001
DEFINITION  Sequence 7 from patent US 6297364.
ACCESSION   AR171866
VERSION     AR171866.1  GI:17910816
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1946)
  AUTHORS   Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
            and Old,L.J.
  TITLE     Isolated nucleic acid molecule encoding cancer associated antigen,
            the antigen itself, and uses thereof
  JOURNAL   Patent: US 6297364-A 7 02-OCT-2001;
FEATURES             Location/Qualifiers
     source          1..1946
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match        97.7%;   Score 1669.2;   DB 6;   Length 1946;
Best Local Similarity   99.8%;   Pred. No. 0;
Matches 1676;   Conservative 0;   Mismatches 3;   Indels 0;   Gaps 0;

Qy    30  CCCGGGGAGCCATCATGAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT  89
          | |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268  CCAGGCAAGCCATCATGAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT  327

Qy    90  CCTACATCCCCGATGAGCAGATAGCACAGGGACCTGAGGACTTCACCTGTGCAGCCAAGCAGC  149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328  CCTACATCCCCGATGAGCAGATAGCACAGGGACCTGAGGACTTCACCTGTGCAGCCAAGCAGC  387

Qy   150  GCTCTCGGGGTCAGCCCCGCCAGGGACCCCGCCAGGGGCCCCAGCCAAGCAGC  209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388  GCTCTCGGGGTCAGCCCCGCCAGGGACCCCGCCAGGGGCCCCAGCCAAGCAGC  447

Qy   210  AGCAAGTGACATCCCCCTTCGCTCCTGGTGCCACCAGTATGTGGTGCCATTATTG  269
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448  AGCAAGTGACATCCCCCTTCGCTCCTGGTGCCACCAGTATGTGGTGCCATTATTG  507

Qy   270  GCAAGGAGGGGGCCACCATCCGCCAACATCACAAAAACGACCCAGTCAAGATAGACGTGC  329
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508  GCAAGGAGGGGGCCACCATCCGCCAACATCACAAAAACGACCCAGTCAAGATAGACGTGC  567

Qy   330  ATAGGAAGGAGAACGCAGTGCAGCTGAAAAAGCCATCAGTCGTGCCACTCCACCCCTGAGG  389
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   568  ATAGGAAGGAGAACGCAGTGCAGCTGAAAAAGCCATCAGTCGTGCCACTCCACCCCTGAGG  627

Qy   390  GCTGCTCTCCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA  449
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628  GCTGCTCTCCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA  687

Qy   450  AAACGGCTGACGAGGTTCCCCTGAAGATCTGCCCCATAATAACTTTGTAGGGCGTCTCA  509
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688  AAACGGCTGACGAGGTTCCCCTGAAGATCTGCCCCATAATAACTTTGTAGGGCGTCTCA  747

Qy   510  TTGGCAAGGAAGGACGGAACCTGAAGAAGGTAGAGAGACAAGAATCCGAGACAAAAATCACCA  569
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748  TTGGCAAGGAAGGACGGAACCTGAAGAAGGTAGAGAGACAAGAATCCGAGACAAAAATCACCA  807

Qy   570  TCTCCCTGCGTTGCAAGAACCTTACCCCTTTACAACCCTGAGAGGACCATCACTCACTG  629
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   808  TCTCCCTGCGTTGCAAGAACCTTACCCCTTTACAACCCTGAGAGGACCATCACTCACTG  867

Qy   630  CCATCGAGAATTGTTGCAGGGGCCGAGCCAGGAATAATGAAAGAAGTTCGGGAGGCCTATG  689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   868  CCATCGAGAATTGTTGCAGGGGCCGAGCCAGGAATAATGAAAGAAGTTCGGGAGGCCTATG  927

Qy   690  AGAATGATGTGGCTGCCATGAGCTCTCACCTGAATCCCGGCCTGAACCTGGCTGCTGTAG  749
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   928  AGAATGATGTGGCTGCCATGAGCTCTCACCTGAATCCCGGCCTGAACCTGGCTGCTGTAG  987

Qy   750  GTCTTTTTCCCAGCTTCATCCAGCGCAGTCCCGCGCCTCCCAGCAGCGTTACTGGGGCTG  809
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   988  GTCTTTTTCCCAGCTTCATCATGAGCGCAGTCCCGCGCCTCCCAGCAGCGTTACTGGGGCTG 1047

Qy   810  CTCCCTATAGTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCG  869
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1048  CTCCCTATAGTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCG 1107

Qy   870  CCCAGGCAGTGGGCGCCATCATCGGCAAGAAGAGGGCAGCACATCAAACAGTCTCCCGGT  929
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1108  CCCAGGCAGTGGGCGCCATCATCGGCAAGAAGAGGGCAGCACATCAAACAGTCTCCCGGT 1167

Qy   930  TTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTCGTATGG  989
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1168  TTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTCGTATGG 1227

Qy   990  TTATCATCACTGGACCCGCCAGAGGCCCAATTCAAGGCTCAGGGAGGAAGAATCTATGGCAAAC 1049
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1228  TTATCATCACTGGACCCGCCAGAGGCCCAATTCAAGGCTCAGGGAGGAAGAATCTATGGCAAAC 1287

Qy  1050  TCAAGGAGGAGAACTTCTTTGTCCAAGGAGGAGAAGTGAAGCTGGAGACCCACATACGTG 1109
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1288  TCAAGGAGGAGAACTTCTTTGTCCAAGGAGGAGAAGTGAAGCTGGAGACCCACATACGTG 1347

Qy  1110  TGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGTGAAAAACGGTGAACGAGTTGC 1169
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1348  TGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGTGAAAAACGGTGAACGAGTTGC 1407

Qy  1170  AGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCACCAGAACCCCTGATGAGAACGACC 1229
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1408  AGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCACCAGAACCCCTGATGAGAACGACC 1467

Qy  1230  AGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGTGGCTCAACGGAAGATCC 1289
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1468  AGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGTGGCTCAACGGAAGATCC 1527

Qy  1290  GAGACATCCTGGCCAGGTTAAGCAGCAGCATCAGAGGACAGAGTAACCAGGCCCAGG 1349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1528  GAGACATCCTGGCCAGGTTAAGCAGCAGCATCAGAGGACAGAGTAACCAGGCCCAGG 1587

Qy  1350  CACGGAGGAAGTGACCAGCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACAACAACGGGACAGAA 1409
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  1588  CACGGAGGAAGTGACCAGCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACAACAACGGGACAGAA 1647

Qy  1410  ATCGGAGAGTGTGCTCTCCCCGCAGGCCCGAGAATGAGTGGGAATCCGGAGCACACACNTGGGC 1469
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1648  ATCGGAGAGTGTGTCTCTCCCCGCAGGCCCGAGAATGAGTGGGAATCCGGAGCACACACNTGGGC 1707

Qy  1470  CGGGCTGTGAGATCAGGTTTGCCCACTGATTGAGAGAAAGATGTTCAGTGAGGAACCCTGA 1529
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1708  CGGGCTGTGAGATCAGGTTTGCCCACTGATTGAGAGAAAGATGTTCAGTGAGGAACCCTGA 1767

Qy  1530  TCTNTCAGCCCCAAACACCCACTTTAAAGATGGCCAAACACTGTNTGCCCTCGGGGTGTCAG 1589
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1768  TCTNTCAGCCCCAAACACCCACTTTAAAGATGGCCAAACACTGTNTGCCCTCGGGGTGTCAG 1827

Qy  1590  AAATTNTAGCGCAAGGCACTTTAAACGTGGAAGAAAATAAAATTTCCTTCAGGTTTTAAAA 1649
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1828  AAATTNTAGCGCAAGGCACTTTAAACGTGGAAGAAAATAAAATTTCCTTCAGGTTTTAAAA 1887

Qy  1650  CAAGAGGGTGGATCACACCCTCAGTGGGAAGAAAAATAAAATTTCCTTCAGGTTTTAAAA 1708
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1888  CAAGAGGGTGGATCACACCCTCAGTGGGAAGAAAAATAAAATTTCCTTCAGGTTTTAAAA 1946

RESULT 5
AR343075
LOCUS       AR343075                1946 bp    DNA    linear    PAT 17-AUG-2003
DEFINITION  Sequence 7 from patent US 6576755.
ACCESSION   AR343075
VERSION     AR343075.1  GI:33738477
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
```

REFERENCE   Unclassified.
            1 (bases 1 to 1946)
AUTHORS   Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
            and Old,L.J.
TITLE   Isolated nucleic acid molecule encoding cancer associated antigen,
            the antigen itself, and uses thereof
JOURNAL   Patent: US 6576756-A 7 10-JUN-2003;
FEATURES   Location/Qualifiers
source   1..1946
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN

Query Match   97.7%; Score 1669.2; DB 6; Length 1946;
Best Local Similarity   99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

Qy   30 CCCGGGGCAGCCATCGAATGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT 89
           |||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   268 CCAGGCAAGCCATCATGAAGTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT 327

Qy   90 CCTACATCCCGATGAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGGCTTTG 149
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 CCTACATCCCGATGAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGGCTTTG 387

Qy   150 GCTCTCGGGGTCAGCCCGCCAGGGCTCACCTGTGCAGCCCAGCCAGCAGC 209
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 GCTCTCGGGGTCAGCCCGCCAGGGCTCACCTGTGCAGCCCAGCCAGCAGC 447

Qy   210 AGCAAGTGGACATCCCCCTTGGCCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTG 269
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448 AGCAAGTGAACATCCCCTTGGCCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTG 507

Qy   270 GCAAGGAGGAGGAGGAGACGGGCCACCATCCGCAACATCACAAAACAGACCCAGTTGGAC 329
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508 GCAAGGAGGAGGAGGAGACGGGCCACCATCCGCAACATCACAAAACAGACCCAGTTGGAC 567

Qy   330 ATAGGAAGGAGAACGCAGTCAGCTGAAGATTATGCATAAAGAGGCTAAGGACACCA 389
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 ATAGGAAGGAGAACGCAGTCAGCTGAAGATTATGCATAAAGAGGCTAAGGACACCA 627

Qy   390 GCTGCTCCTCCGCTTGTAAGATGACTTGGAGATTCCACCCCTCAGCCCCTGAGG 449
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628 GCTGCTCCTCCGCTTGTAAGATGACTTGGAGATTCCACCCCTCAGCCCCTGAGG 687

Qy   450 AAACGGCTGACGAGGTTCCCTGAGAGTCTGGCCCATAATAACTTGTAGGGCGGTCTCA 509
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688 AAACGGCTGACGAGGTTCCCTGAGAGTCTGGCCCATAATAACTTGTAGGGCGGTCTCA 747

Qy   510 TTGGCAAGGAGGACGGAACCTGAAGGAGTAGAGCAAGTAGAGCAAGAAAAATCACCA 569
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748 TTGGCAAGGAGGACGGAACCTGAAGGAGTAGAGCAAGTAGAGCAAGAAAAATCACCA 807

Qy   570 TCTCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGGGG 629
           ||||||||||||||||||||||||||||||||||||||||||||
Db   808 TCTCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGGGG 867

Qy   630 CCATCGAGAATTGTTGCAGGGCCGAGCAGGAGAAAGTTCGGGAGGCCTATG 689
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   868 CCATCGAGAATTGTTGCAGGGCCGAGCAGGAGAAAGTTCGGGAGGCCTATG 927

Qy   690 AGAATGATGTGGCTGCCATGAGCTCTCACCTGGCTGAACCTGGCTGTGTAG 749
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   928 AGAATGATGTGGCTGCCATGAGCTCTCACCTGGCTGAACCTGGCTGTGTAG 987

Qy   750 GTCTTTTCCAGCTTCATCCAGCGAGTCCCGCCTCCCGCGCCGCCGGGGCTG 809
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   988 GTCTTTTCCAGCTTCATCCAGCGAGTCCCGCCTCCCGCGCCGCCGGGGCTG 1047

Qy   810 CTCCCTATAGCTCCTTTATGCAGGCTCCCCAGCGCAGGAGATAGTGTCAGG 869
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1048 CTCCCTATAGCTCCTTTATGCAGGCTCCCCAGCGCAGGAGATAGTGTCAGG 1107

Qy   870 CCCAGGCAGTGGGGCCGGCCCATCATCGGCAGAAGGGGCAGCAGCATCAAACAGCTCCCCGGT 929
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1108 CCAGGCAGTGGGGCCGGCCCATCATCGGCAGAAGGGGCAGCCACATCAAACAGCTCCCCGGT 1167

Qy   930 TTGCCAGCGCCTCCATCAAGATTGCACCACCGAAACACCTGACTCCAAAGTTCGTATGG 989
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1168 TTGCCAGCGCCTCCATCAAGATTGCACCACCGAAACACCTGACTCCAAAGTTCGTATGG 1227

Qy   990 TTATCATCACTGGACCGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAAC 1049
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1228 TTATCATCACTGGACCGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAAC 1287

Qy   1050 TCAAGGAGGAGAACTTCTTTGGTCCCAGGAGGAGAGTGAAGCTGGAGAGCCACATACGTG 1109
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1288 TCAAGGAGGAGAACTTCTTTGGTCCCAGGAGGAGAGTGAAGCTGGAGAGCCCACATACGTG 1347

Qy   1110 TGCCAGCATCAGAGCTGCGCCGGTCATTGGCAAAGTGGAAAAACGGTGAACGAGTTGC 1169
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1348 TGCCAGCATCAGAGCTGCGCCGGTCATTGGCAAAGTGGAAAAACGGTGAACGAGTTGC 1407

Qy   1170 AGAATTGACGGCAGCTGAGGTGTAGTACCAAGAGACCAGACCCCTGATGAGAACGACC 1229
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1408 AGAATTGACGGCAGCTGAGGTGTAGTACCAAGAGACCAGACCCCTGATGAGAACGACC 1467

Qy   1230 AGGTCATCGTGAAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC 1289
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1468 AGGTCATCGTGAAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC 1527

Qy   1290 GAGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAACGGACAGAGTAACCAGGCCCAGG 1349
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1528 GAGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAACGGACAGAGTAACCAGGCCCAGG 1587

Qy   1350 CACGGAGGAAGTGACCAGCAGCCCCTCCCTNGAGGCCTGAGATAGACGTGC 1409
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1588 CACGGAGGAAGTGACCAGCAGCCCCTCCCTGTTCCCTNGAGTCTCAGGCAGAA 1647

Qy   1410 ATCGAGAGTGTGCTCTCCCCGGCCTGAGAATGAGTGGAATCCGGGACACNTGGGC 1469
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1648 ATCGAGAGTGTGCTCTCCCCGGCCTGAGAATGAGTGGAATCCGGGACACNTGGGC 1707

Qy   1470 CGGGCTGTAGATCAGGTTTGCCCACTCAGTTGGAAAGATGTTCCAGTGAGGAACCCTGA 1529
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1708 CGGGCTGTAGATCAGGTTTGCCCACTCAGTTGGAAAGATGTTCCAGTGAGGAACCCTGA 1767

Qy   1530 TCTNTCAGCCCAAACACCCACCAATTGGCCCACACACTGTNTGCCCCCTCGGGGTGTCAG 1589
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1768 TCTNTCAGCCCAAACACCCACCAATTGGCCCACACACTGTNTGCCCCCTCGGGGTGTCAG 1827

Qy   1590 AAATTNTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCAC 1649
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1828 AAATTNTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCAC 1887

Qy   1650 CAAGAGGGTGGATCACACCTCAGTGGGAAGAGAAAAATAAAATTCCTTCAGGTTTAAAA 1708
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1888 CAAGAGGGTGGATCACACCTCAGTGGGAAGAGAAAAATAAAATTCCTTCAGGTTTAAAA 1946

RESULT 6
BD209926

JOURNAL   antigens, these antigens and method of using the same
           Patent: JP 200251204049-A 5 23-APR-2002;
           LUDWIG INSTITUTE FOR CANCER RESEARCH
COMMENT   OS   Homo sapiens (human)
           PN   JP 200251204049-A/5
           PD   23-APR-2002
           PF   16-MAR-1999 JP 2000545030
           PR   17-APR-1998 US   09/061709
           PI   YAO TSENG CHEN,ALI GURE,SOLAM TSANG,ELISABETH STOCKERT,ELKE
           PI   JAGER,
           PI   ALEXANDER KNUTH,LLOYD J OLD
           PC   C12N15/09,A61K35/12,A61K39/00,A61K39/39,A61P35/00,C07K16/32,
           PC   C12N15/00,A61K39/00,
           PC   C12N1/15,
           PC   C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/ PC
           574,
           PC   C12N15/00,C12N5/00
           CC   n=not determined(1622,1702,1771,1809,1833)
FH   Key   Location/Qualifiers
FT   source   1..1946
FT   /organism='Homo sapiens (human)'.
FEATURES   Location/Qualifiers
           source   1..1946
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   97.7%;   Score 1669.2;   DB 6;   Length 1946;
Best Local Similarity 99.8%;   Pred. No. 0;
Matches 1676;   Conservative 0;   Mismatches 3;   Indels 0;   Gaps 0;

Qy   30 CCCGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT 89
           ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT 327

Qy   90 CCTACATCCCCGATGAGCAGATAGCACAGGGACCTGAGAAGAATGGGCGCGCCGAGGGCTTTG 149
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 CCTACATCCCCGATGAGCAGATAGCACAGGGACCTGAGAAGAATGGGCGCGCCGAGGGGCTTTG 387

Qy   150 GCTCTCGGGTCAGCCCCGCCAGGGCTCACCTGTGGCCAGCGGGCGGGGGCCCAGCCAAGCAGC 209
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 GCTCTCGGGGTCAGCCCCGCCAGGGGCTCACCTGTGGCCAGCGGGCGGGGGCCCAGCCAAGCAGC 447

Qy   210 AGCAAGTGGACATCCCCCTTCGGCTCTGTGCCCACCCAGTATGTGGGTGCCATTATTG 269
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448 AGCAAGTGGACATCCCCCTTCGGCTCTGTGCCCACCCAGTATGTGGGTGCCATTATTG 507

Qy   270 GCAAGGAGGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGC 329
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508 GCAAGGAGGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGC 567

Qy   330 ATAGGAAGGAGACGCAGTGCAGCTGCACCTGTGAAAAGCCATCAGTGTGCACTCACCCCTGAGG 389
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 ATAGGAAGGAGACGCAGTGCAGCTGCACCTGTGAAAAGCCATCAGTGTGCACTCACCCCTGAGG 627

Qy   390 GCTGCTCTCCGCGCTTGTAAGATGATCTTGGAGATCCTGAAGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 449
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628 GCTGCTCTCCGCGCTTGTAAGATGATCTTGGAGATCCTGAAGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 687

Qy   450 AAACGGCTGACGAGGTTCCCTGAAGATCCTGAAGATCTTGGAGATCCTGAAGATCTTGGAGATTATGTAGGGCGTCTCA 509
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688 AAACGGCTGACGAGGTTCCCTGAAGATCCTGAAGATCTTGGAGATTATGTAGGGCGTCTCA 747

Qy   510 TTGGCAAGGAACGGAACCTGAACTGAAGGTAGACCAAGATACCGAGACAAAAATCACCA 569
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748 TTGGCAAGGAACGGAACCTGAACTGAAGGTAGACCAAGATACCGAGACAAAAATCACCA 807

Qy   570 TCTCCTCGTTGCAAGACCTTACAACCCTTACAACCCTTACAACCCTTACAACCCTGAGAGGAGGAC 629
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   808 TCTCCTCGTTGCAAGACCTTACAACCCTTACAACCCTTACAACCCTGAGAGGAGGAC 867

Qy   630 CCATCGAGAATTGTTGCAGGGGCCGAGCGGCAGCAGGAGAAATAATGAAGAAGAGTTCGGGAGGCCTATG 689

868 CCATCGAGAATTGTTGCGGCTGCCATGAGCTCTCACCTGATCCTGGCCTGGCCTGCTGCTGTAGG
    Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 927

Qy   690 AGAATGATGTGGCTGCTGCCATGAGCTCTCACCTGATCCTGGCCTGGCCTGCTGCTGTAG 749
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   928 AGAATGATGTGGCTGCTGCCATGAGCTCTCACCTGATCCTGGCCTGGCCTGCTGCTGTAG 987

Qy   750 GTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCGCCTCCCAGCAGCGTTACTGGGGCTG 809
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   988 GTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCGCCTCCCAGCAGCGTTACTGGGGCTG 1047

Qy   810 CTCCCATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCG 869
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1048 CTCCCATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCG 1107

Qy   870 CCCAGGCAGTGGGCGCCATCATCGGCAAGAGAGGGCAGCACATCAAACAGCTCTCCCGGT 929
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1108 CCCAGGCAGTGGGCGCCATCATCGGCAAGAGAGGGCAGCACATCAAACAGCTCTCCCGGT 1167

Qy   930 TTGCCAGCGCCTCCATCAAGATGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCAAGGCTCATGGCAAAC 989
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1168 TTGCCAGCGCCTCCATCAAGATTGCACCACCGAAACACCTGACTCGACTCCAAAGTTCGTATGG 1227

Qy   990 TTATCATCACTGGACCGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCAAGGCTCATGGCAAAC 1049
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1228 TTATCATCACTGGACCGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCAAGGCTCATGGCAAAC 1287

Qy   1050 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAAGTGAAGCTGGAGCCCACATACGTG 1109
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1288 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAAGTGAAGCTGGAGACCCACATACGTG 1347

Qy   1110 TGCCAGCATCAGCAGCTGGGTCATTGGCAAAAACGGTGAACGAGTTGC 1169
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1348 TGCCAGCATCAGCAGCTGGGTCATTGGCAAAAACGGTGAACGAGTTGC 1407

Qy   1170 AGAATTGACGCGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGAACGACC 1229
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1408 AGAATTGACGCGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGAACGACC 1467

Qy   1230 AGGTCATCGTGAAAATCATCGGACATTCATCATGCCAGTCAGATGGCTCAACGGAAGATCC 1289
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1468 AGGTCATCGTGAAAATCATCGGACATTTCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC 1527

Qy   1290 GAGACATCCTGCCAAGGTTAAGCAGCATCAGAAGGGACAGAGGACAGAGTAACCAGGCCCAGG 1349
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1528 GAGACATCCTGCCAAGGTTAAGCAGCATCAGAAGGGACAGAGGACAGAGTAACCAGGCCCAGG 1587

Qy   1350 CACGGAGGAAGTGACCAGCCCCTCCCTGTCCCTTNGAGTCCAGAACAACACGGGCAGAA 1409
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1588 CACGGAGGAAGTGACCAGCCCCTCCCTGTCCCTTNGAGTCCAGAACAACACGGGCAGAA 1647

Qy   1410 ATCGAGAGTGTGCTTGCTCTCCCCCGCAGGCCTGAGAATGAGTGGGAATCCGGACACNTGGGC 1469
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1648 ATCGAGAGTGTGCTTGCTCTCCCCCGCAGGCCTGAGAATGAGTGGGAATGGGAATCCGGACACNTGGGC 1707

Qy   1470 CGGGCTGTAGATCAGGTTGCCCACTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGA 1529
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1708 CGGGCTGTAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGA 1767

Qy   1530 TCTNTCAGCCCCAAACACCCAATTGGCCCAATTGGCCCAACACTGTNTGCCCCTNGCCGGGGTGTCAG 1589
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1768 TCTNTCAGCCCCAAACACCCAATTGGCCCAACACTGTNTGCCCCTNGCCGGGGTGTCAG 1827

Qy   1590 AAATTNTAGGCCAAGGCACTTTAAACGTGGATTGTTAAAGAAGCTCTCCAGGCCCCAC 1649
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1828 AAATTNTAGGCCAAGGCACTTTAAACGTGGATTGTTAAAGAAGCTCTCCAGGCCCCAC 1887

Qy   1650 CAAGAGGGGTGATCACACCTCAGTGGGAAGAAAAATAAAATTTCCTTCAGGTTTTAAA 1708
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1888 CAAGAGGGGTGATCACACCTCAGTGGGAAGAAGAAATAAAATTTCCTTCAGGTTTTAAA 1946

Qy   1630 CCATCGAGAATTGTTGCAGGGGCCGAGCGGCAGCAGGAGAAATAATGAAGAAGAGTTCGGGAGGCCTATG

RESULT 7

```
LOCUS       AF117106              2130 bp    mRNA    linear   PRI 26-JAN-1999
DEFINITION  Homo sapiens IGF-II mRNA-binding protein 1 (IMP-1) mRNA, complete
            cds.
ACCESSION   AF117106
VERSION     AF117106.1  GI:4191607
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2130)
  AUTHORS   Nielsen,J., Christiansen,J., Lykke-Andersen,J., Johnsen,A.H.,
            Wewer,U.M. and Nielsen,F.C.
  TITLE     A family of insulin-like growth factor II mRNA-binding proteins
            represses translation in late development
  JOURNAL   Mol. Cell. Biol. 19 (2), 1262-1270 (1999)
   MEDLINE   99108099
    PUBMED   9891060
REFERENCE   2  (bases 1 to 2130)
  AUTHORS   Nielsen,J., Christiansen,J., Lykke-Andersen,J., Johnsen,A.H.,
            Wewer,U.M. and Nielsen,F.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1998) Institute of Molecular Biology, University
            of Copenhagen, Soelvgade 83H, Copenhagen DK-1307, Denmark
FEATURES             Location/Qualifiers
     source          1..2130
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     gene            1..2130
                     /gene="IMP-1"
     CDS             10..1743
                     /gene="IMP-1"
                     /note="binds multiple sites in IGF-II mRNA 5'UTR"
                     /codon_start=1
                     /product="IGF-II mRNA-binding protein 1"
                     /protein_id="AAD09826.1"
                     /db_xref="GI:4191608"
                     /translation="MNKLYIGNLNESVTPADLEKVFAEHKISYSGQFLVKSGYAFVDC
                     PDEHWAMKAIETFSGKVELQGKRLEIEHSVPKKQRSRKIQIRNIPPQLRWEVLDSLLA
                     QYGTVENCEQVNTESETAVNVTYSNREQTRQAIMKLNGHQLENHALKVSYIPDEQIA
                     QGPENGRRGGFGSRGQPRQGSPVAAGAPAKQQQVDIPLRLLVPTQYVGAIIGKEGATI
                     RNITKQTGSKIDVHRKENAGAAEKAISVHSTPEGCSSACKMIEIMHKEAKDTKTADE
                     VPLKILAHNNFVGRLIGKEGRNLKKVEQDTETKITISSLQDLTLYNPERTITVKGAIE
                     NCCRAEQEIMKKVREAYENDVAAMSLQSHLIQGLNLAAVGLFPASSAVPPPSSVTG
                     AAPYSSFMQAPEQEMVQVFIPAQAVGAIIGKKGQHIKQLSRFASASIKIAPPETPDSK
                     VRMVIITGPPEAQFKAQGRIYGKLKEENFPGPKEEVKLETHIRVPASAAGRVIGKGGK
                     TVNELQNLTAAEVVVPRDQTPDENDQVIVKIIGHFYASQMAQRKIRDILAQVKQQHQK
                     GQSNQAQRRK"
ORIGIN
```

Query Match   96.8%;   Score 1653.2;   DB 9;   Length 2130;
Best Local Similarity   99.2%;   Pred. No. 0;
Matches 1671;   Conservative 0;   Mismatches 8;   Indels 6;   Gaps 1;

```
Qy        30 CCGGGGGAGCCATCATGAAGTGAATGGCCACCAGTTGAGAACCATGCCTGAAGGTCT  89
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       404 CCAGGCAAGCCATCATGAAGTGAATGGCCACCAGTTGAGAACCATGCCTGAAGGTCT  463

Qy        90 CCTACATCCCCGATGAGCAGATAGCAGCAGGGACCTGAGAATGGGGCGGGGGGCTTTG  149
             ||||| || |||||||||||||||||||||||||||||||||||||||||||||||||
Db       464 CCTACATCCCCGATGAGCAGATAGCAGCAGGGACCTGAGAATGGGCGGGGGGCTTTG  523

Qy       150 GCTCTCGGGGTCAGCCCCAGGGCTCACCTGTGCCACCAGTTGGGCTCACCTGTGCCACCA  209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       524 GCTCTCGGGGTCAGCCCCAGGGCTCACCTGTGCCACCAGTTGGGCTCACCTGTGCCACCA  583

Qy       210 AGCAAGTGGACATCCCCCCTTGGCTCCTCCCCTGTGTGCCACCAGTTATGTGGGGTGCCATTATTG  269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       584 AGCAAGTGGACATCCCCCCTTGGCTCCTCCCCTGTGTGCCACCAGTTATGTGGGGTGCCATTATTG  643

Qy       270 GCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCCAGTCCAAGATAGACGTGC  329
```

Db   644 GCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCCAGTCCAAGATAGACGTGC 703

Qy   330 ATAGGAAGGAAGCAGGTGCAGCTGCAGCTGCAGTTGCACTCAGTGTGCACTCCACCCTGAGG 389
Db   704 ATAGGAAGGAAGCAGGTGCAGCTGCAGCTGCAGTTGCACTCAGTGTGCACTCCACCCTGAGG 763

Qy   390 GCTGCTCCTCCGCTGTAAGATGATCTTGGAGATTATGCATAAAGAGAGCTAAGGACACCA 449
Db   764 GCTGCTCCTCCGCTGTAAGATGATCTTGGAGATTATGCATAAAGAGAGCTAAGGACACCA 823

Qy   450 AAACGCTGACGAGGTTCCCCTGAAGATCCTGGCCCATATAATAACTTTGTAGGGGGTCTCA 509
Db   824 AAACGCTGACGAGGTTCCCCTGAAGATCCTGGCCCATATAATAACTTTGTAGGGGCGTCTCA 883

Qy   510 TTGGCAAGGAAGGACGGAACCTGAAGAAGTGAGCAAGTAGAGCAAGATACCGAGACAAAAAATCACCA 569
Db   884 TTTGGCAAGGAAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAAAATCACCA 943

Qy   570 TCTCCTCGTTGCAAGACCTTACCCTTACAACCCTGAGAGGACCATCACTGTGAAGGGGG 629
Db   944 TCTCCTCGTTGCAAGACCTTACCCTTACAACCCTGAGAGGACCATCACTGTGAAGGGGG 1003

Qy   630 CCATCGAGAATTGTTGCAGGGCCGAGCAGCAGAAATAATGAAGAAGTTCGGGAGGCCTATG 689
Db   1004 CCATCGAGAATTGTTGCAGGGCCGAGCAGCAGAAATAATGAAGAAGTTCGGGAGGCCTATG 1063

Qy   690 AGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCTGAACCTGGCTG 743
Db   1064 AGAATGATGTGGCTGCCATGAGCCCTGATCTCTCACCTGATCCCTGAACCTGGCTG 1123

Qy   744 CTGTAGGTCTTTTCCCAGCTTCATCCAGCTTTCATCCAGCGCAGTCCCGCCTTCCCAGCAGCGGTTACTG 803
Db   1124 CTGTAGGTCTTTTCCCAGCTTCATCCAGCTTTCATCCAGCGCAGTCCCGCCTTCCCAGCAGCGGTTACTG 1183

Qy   804 GGGCTGCTCCCCTATAGTCTCCTTTATGCAGCTCCCGAGCAGGGGTGCAGCACATCAAACAGTCT 863
Db   1184 GGGCTGCTCCCCTATAGTCTCCTTTATGCAGCTCCCGAGCAGGGGTGCAGCACATCAAACAGTCT 1243

Qy   864 TCCCCGCCAGCAGTGGGCCCATCAATGGCAAGAACGGGGCAGCACATCAAACAGCTCT 923
Db   1244 TCCCCGCCAGCAGTGGGCCCATCAATCGCAAGAACGGGGCAGCACATCAAACAGCTCT 1303

Qy   924 CCCGGTTTGCCAGCGCCTCATCAAGATTGCACACCGCAAACACCTGACTCCAAAGTTC 983
Db   1304 CCCGGTTTGCCAGCGCCTCATCAAGATTGCACACCGCAAACACCTGACTCCAAAGTTC 1363

Qy   984 GTATTGGTTATCATCACTGACCCGCAGAGGCCCAATTCAAGGCTCAGGGAGAATCATG 1043
Db   1364 GTATTGGTTATCATCACTGACCCGCAGAGGCCCAATTCAAGGCTCAGGGAGAATCATG 1423

Qy   1044 GCAAACTCAAGGAGGAGGAGAACTTCTTTGTCCCAAGGAGGAAGTGGAGACCCACA 1103
Db   1424 GCAAACTCAAGGAGGAGGAGAACTTCTTTTGTCCCGGGTCATTGGCAAGGTGGAAAACG 1483

Qy   1104 TACGTGTGCCAGCATCAGCAGCTCGGCCGGTCATTGGCAAGGTGGAAAAACGTGAACG 1163
Db   1484 TACGTGTGCCAGCATCAGCAGCTCGGCCGGTCATTGGCAAGGTGGAAAAACGTGAACG 1543

Qy   1164 AGTTGCAGAATTTGACGGACAGAGTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGA 1223
Db   1544 AGTTGCAGAATTTGACGGACAGAGTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGA 1603

Qy   1224 ACGACCAGTCATCGTGAAAATCATCGGACATTTCATGCCCAGTCACAAGCAGC 1283
Db   1604 ACGACCAGTCATCGTGAAAATCATCGGACATTTCATGCCCAGTCACAAGCAGC 1663

Qy   1284 AGATCCGAGACATCCTGGCCCAGTTAAGCACCAGCATCAGAGGGACAGAGTAACCAGG 1343
Db   1664 AGATCCGAGACATCCTGGCCCAGTTAAGCACCAGCATCAGAGGGACAGAGTAACCAGG 1723

Qy   1344 CCCAGGCACGGGAGAAGTGAACAGCCCTCCCTTGCCCTGTCCCTTNGAGTCAGGACAACAACGG 1403

```
VPLKTLAHNFVGRLIGKEGRNLKKVEQDTETKITISSLQDLTLYNPRTITVKGAIE
NCCRAEQEIMKKVREAYENDVAAMSLQSHLTPGLNLAAVGLFPASSSAVPPPSSVTG
AAPYSSFMOAPEQEMVQVPIPAQVGVAIIGKKQHIKQLSRFASASIKIAPPETPDSK
VRMVIITGPPEAQFKAQGRIYGKLKEENFPGPKEEVKLETHIRVPASAAGRVIGKGGK
TVNELQNLTAAEVVVPRDQTPDENDQVIVKIIGHFYASQMAQRKIRDILAQVKQQHQK
GQSNQAQARRK"
```

ORIGIN

```
Query Match           95.5%;   Score 1631.8;   DB 9;   Length 2381;
Best Local Similarity 98.9%;   Pred. No. 0;
Matches 1663;         Conservative 0;   Mismatches 0;   Indels 12;   Gaps 2;
```

```
Qy    30  CCCGGGGAGCCATCATGAAGTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT   89
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   700  CCAGGCAAGCCATCATGAAGTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT  759

Qy    90  CCTACATCCCGATGAGCAGATAGCACAGGAGACCTGAGCAGTGGGCGGCCGAGGGGCTTTG  149
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   760  CCTACATCCCGATGAGCAGATAGCACAGGAGACCTGAGCAGTGGGCGGCCGAGGGGCTTTG  819

Qy   150  GCTCTCGGGGTCAGCCCCGCCCAGGGCTCACCTGTGGCCAGCCAAGCAGC          209
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   820  GCTCTCGGGGTCAGCCCCGCCCAGGGCTCACCTGTGGCCAGCCAAGCAGC          879

Qy   210  AGCAAGTGGACATCCCCCTTCGGCTTTCGGCTTCCCCAGTATGTGGTGCCCATTATTG     269
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   880  AGCAAGTGGACATCCCCCTTCGGCTTTCGGCTTCCCCAGTATGTGGTGCCCATTATTG     939

Qy   270  GCAAGGAGGGGGCCACCATCCGCAACATCACAAAGCAGACACCCAGTCCAAGATAGACGTGC  329
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   940  GCAAGGAGGGGGCCACCATCCGCAACATCACAAAGCAGACACCCAGTCCAAGATAGACGTGC  999

Qy   330  ATAGGAAGGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCAGTGTGCACTCCGAGG  389
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1000  ATAGGAAGGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCAGTGTGCACTCCGAGG  1059

Qy   390  GCTGCTCCTCGGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA    449
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1060  GCTGCTCCTCGGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA   1119

Qy   450  AAACGGCTGACGAGTTCCCTGAAGATCCTGCGCCATAATAACTTTGTAGGGCGTCTCA     509
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1120  AAACGGCTGACGAGTTCCCTGAAGATCCTGCGCCATAATAACTTTGTAGGGCGTCTCA    1179

Qy   510  TTGGCAAGGAAGGACGGAACCTGAAGAAGTAGAGCAAGTAGACCAAGAAAAATCACCA      569
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1180  TTGGCAAGGAAGGACGGAACCTGAAGAAGTAGAGCAAGTAGACCAAGAAAAATCACCA     1239

Qy   570  TCTCCTGTTGCAAGACCTTACCCTTTACAACCCTTACAACCCTGAGAGGACCATCACTGTGAAGGGGG  629
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1240  TCTCCTGTTGCAAGACCTTACCCTTTACAACCCTTACAACCCTGAGAGGACCATCACTGTGAAGGGGG  1299

Qy   630  CCATCGAGAATTGTTCAGGGCCGAGCAGGAGAATAATGAAGAAAGTTCGGGAGGCCTATG  689
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1300  CCATCGAGAATTGTTCAGGGCCGAGCAGGAGAATAATGAAGAAAGTTCGGGAGGCCTATG  1359

Qy   690  AGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCTCGGCCTGCCTCCCGCC  743
          |||||||||||||||||||||||      |||||||||||||||||||||||||||||
Db  1360  AGAATGATGTGGCTGCCATGAGCTCTCACCTGATCCTCGGCCTGCCTCCCGCC         1419

Qy   744  CTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCTCCCAGCAGCGGTTACTG     803
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1420  CTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCTCCCAGCAGCGGTTACTG    1479

Qy   804  GGGCTGCTCCCTATAGCTCCTTTATGCAGCGTCAGGTGTCAGGTGTTTA            863
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1480  GGGCTGCTCCCTATAGCTCCTTTATGCAGCGTCAGGTGTCAGGTGTTTA            1539

Qy   864  TCCCCGCCCAGCAGTGGGCGGCCATCATCGGCAAGGAGGGCGGCAGCACATCAAACAGCTCT  923
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1540  TCCCCGCCCAGCAGTGGGCGGCCATCATCGGCAAGGAGGGCGGCAGCACATCAAACAGCTCT  1599

Qy   924  CCCGGTTTGCCAGCGCCTCCATCAAGATTGCCAAGATTGCCACACCCGAAACACCTGACTCCAAGTTC  983
```

```
Db  1724  CCCAGGCACGGAGGAAGTGACCAGCCCCTCCCTGTCCCTTCGAGTCCAGGACAACAACGG  1783
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1404  GCAGAAATCAGGAGAGTGTGCTCTCCCCGGCAGGCCTGAGAATGAGTGGGAATCCGGACAC  1463
          |||||||||| ||||||||||| ||||||||||||||||||||||||||||
Db  1784  GCAGAAATCGAGAGTGTGTCTCTCCCCCGGCAGGCCTGAGAATGAGTGGGAATCCGGACAC  1843

Qy  1464  NTGGGGCCGGGCTGTAGATCAGGTTGCCCACTGATTGAGAAAGATGTTCCAGTGAGGAA  1523
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1844  CTGGGGCCGGGCTGTAGATCAGGTTGCCCACTGATTGAGAAAGATGTTCCAGTGAGGAA  1903

Qy  1524  CCCTGATCNTCAGCCCCAAACACCCAACACCCACCCCAATTGGCCAACACTGTNTGCCCCTCGGGG  1583
          |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  1904  CCCTGATCTCTCAGCCCCAAACACCCAACACCCACCCCAATTGGCCAACACTGTCTGCCCCTCGGGG  1963

Qy  1584  TGTCAGAAATTNTAGCGCAAGGACTTTAAACGTGGATTGTTTAAAGAAGTCTCCAGG  1643
          |||||||||| |||||||||||||||||||||||||||||||||||||||
Db  1964  TGTCAGAAATTCTAGCGCAAGGACTTTAAACGTGGATTGTTTAAAGAAGTCTCCAGG  2023

Qy  1644  CCCACCAAGAGAGGGTGATCACACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGTTT  1703
          ||||||||||||||||||||||||||||||||||||||||| ||||||||
Db  2024  CCCACCAAGAGAGGGTGATCACACCTCAGTGGGAAGAAGAAAATAAAATTTCCTTCAGGTTT  2083

Qy  1704  TAAAA  1708
          |||||
Db  2084  TAAAA  2088
```

RESULT 8
AF198254

```
LOCUS       AF198254                2381 bp    mRNA    linear   PRI 02-MAR-2000
DEFINITION  Homo sapiens mRNA-binding protein CRDBP mRNA, complete cds.
ACCESSION   AF198254
VERSION     AF198254.1  GI:7141071
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2381)
  AUTHORS   Ioannidis,P., Trangas,T., Dimitriadis,E., Samiotaki,M.,
            Panoutsakopoulos,G., Kyriazoglou,I., Voutzoulias,S.,
            Tsiapalis,C.M., Kittas,C., Agnantis,N. and Pandis,N.
  TITLE     Ectopic expression of a KH-domain containing protein, highly
            homologous to both human IMP-1 and mouse CRD-BP, in benign and
            malignant mesenchymal tumors
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2381)
  AUTHORS   Ioannidis,P., Trangas,T., Dimitriadis,E. and Samiotaki,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-OCT-1999) G. Papanikolaou Research Center, St Savvas
            Hospital, 171 Alexandras Avenue, Athens, Attica 11522, Greece
FEATURES             Location/Qualifiers
     source          1..2381
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17q21.3-q22"
                     /cell_line="MCF-7"
     CDS             306..2039
                     /note="KH-domain containing protein; similar to Mus
                     musculus CRD-BP (c-myc mRNA coding region
                     determinant-binding protein) and Homo sapiens IMP-1
                     (insulin growth factor II mRNA-binding protein)"
                     /codon_start=1
                     /product="mRNA-binding protein CRDBP"
                     /protein_id="AAF37203.1"
                     /db_xref="GI:7141072"
                     /translation="MNKLYIGNLNESVTPADLEKVFAEHKISYSGQFLVKSGYAFVDC
                     PDEHWAMKAIETFSGKVELQGKRLEIIHSVPKKQRSRKIQIRNIPPQLRWEVLDSLLA
                     QYGTVENCEQVNTESETAVVNVTYSNREQTRQAIMKLNGHQLENHALKVSYIPDEQIA
                     QGPENGRRGGFGSRGQPRQGSPVAAGAPAKQQVDIPLRLLVPTQYVGAIIGKEGATI
                     RNITKQTQSKIDVHRKENAGAABKAISVHSTPEGCSSACKMLEIMHKEAKDTKTADE"
```

COMMENT

```
HELIX RESEARCH INSTITUTE
OS   Homo sapiens (human)
PN   JP 2002017375-A/3242
PD   22-JAN-2002
PF   07-JUL-2000 JP 2000253172
PI   TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI   ISHII,
PI   YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI   TETSUJI OTSUKI,HISASHI KOGA
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC   C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH   Key
Location/Qualifiers
```

FEATURES
```
FT   CDS           (112)..(1068).
source            (112)..(1068)
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
```

ORIGIN

```
Query Match        81.3%;   Score 1389;   DB 6;   Length 2780;
Best Local Similarity   99.2%;   Pred. No. 0;
Matches 1405;   Conservative   0;   Mismatches   5;   Indels   6;   Gaps   1;

Qy     299 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAACGGAGGAACGCAGGTGCAGCTGAA 358
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAACGGAGGAACGCAGGTGCAGCTGAA 60

Qy     359 AAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCGCTTGTAAGATGATCTTG 418
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCGCTTGTAAGATGATCTTG 120

Qy     419 GAGATTATGCATAAAGAGGCTAAGGACACCAAAAACGGCTGACGAGGTTCCCCTGAAGATC 478
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GAGATTATGCATAAAGAGGCTAAGGACACCAAAAACGGCTGACGAGGTTCCCCTGAAGATC 180

Qy     479 CTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGAAGGAAGGACGAACCTGAAGAAG 538
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGAAGGAAGGACGAACCTGAAGAAG 240

Qy     539 GTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTTTAC 598
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTTTAC 300

Qy     599 AACCCTGAGGAGGACCATCACTGTGAAGGGGGCCATGAGGAATTGTTGCAGGGCCGAGCAG 658
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AACCCTGAGGAGGACCATCACTGTGAAGGGGGCCATGAGGAATTGTTGCAGGGCCGAGCAG 360

Qy     659 GAAATAATGAAGAAAGTTCGGAGGCCTATGAGAAGATGATGGCGTGCTGCCATGAGC----- 712
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GAAATAATGAAGAAAGTTCGGAGGCCTATGAGAAGATGATGGCGTGCTGCCATGAGCGTGCAG 420

Qy     713 TCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGCTTCGGGCGTTCCCAGCTTCATCCAGC 772
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGCTTCGGGCGTTCCCAGCTTCATCCAGC 480

Qy     773 GCAGTCCCGCCGCCTCCCAGCAGCTCCCAGCGTACTGGCGTCTCCTATAGCTCCTTTATGCAG 832
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GCAGTCCCGCCGCCTCCCAGCAGCTCCCAGCGTACTGGCGTCTCCTATAGCTCCTTTATGCAG 540

Qy     833 GCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCAGCAGTGGGCGCCATCATC 892
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCAGCAGTGGGCGCCATCATC 600

Qy     893 GGCAAGAGGGGCAGCACATCAAACAGCTCTCCCGGTTGCCAGCGCCTCATCAAGATT 952
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GGCAAGAGGGGCAGCACATCAAACAGCTCTCCCGGTTGCCAGCGCCTCATCAAGATT 660

Qy     953 GCACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCGCAGAG 1012
```

```
Db    1600 CCCGGTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTTCCAAAGTTC 1659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     984 GTATGGTTATCATCACTGGACCGGCCAGAGGCCCAGGCCCAATTCAAGGCTCAGGGAAGAATCTATG 1043
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1660 GTATGGTTATCATCACTGGACCGGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATG 1719
Qy    1044 GCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAGCTGAGACCCACA 1103
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1720 GCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAGCTGGAGACCCACA 1779
Qy    1104 TACGTGTGCCAGCATCAGCAGCGTGGCCGGGTCATTGGCAAGGTGGAAAAACGGTGAACG 1163
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1780 TACGTGTGCCAGCATCAGCAGCGTGGCCGGGTCATTGGCAAAGGTGGAAAAACGGTGAACG 1839
Qy    1164 AGTTGCAGAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGACACAGACCCCTGATGAGA 1223
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1840 AGTTGCAGAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGACCAGACCCCTGATGAGA 1899
Qy    1224 ACGACCAGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGA 1283
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1900 ACGACCAGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGA 1959
Qy    1284 AGATCCGAGACATCCTGCCCAGGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACCAGG 1343
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1960 AGATCCGAGACATCCTGCCCAGGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACCAGG 2019
Qy    1344 CCCAGGACGGGAGTGACCAGCCCCTCCCTGTCCCTTNGAGTCCAGGACAACAACGG 1403
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2020 CCCAGGCACGAGGAAGTGACCAGCCCCTCCCTGTCCCTTCGAGTCCTGAGGACAACAACGG 2079
Qy    1404 GCAGAAATCGAGAGTGTGCTCTCCCCCGGCCAGGCCCTGAGAGTGAGTGGGAATCCGGGACAC 1463
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2080 GCAGAAATCGAGAGTGTGCTCTCCCCCGGCCAGGCCCTGAGAGTGAGTGGGAATCCGGGACAC 2139
Qy    1464 NTGGGCCCGGCTGTAGATCAGGTTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAA 1523
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2140 CTGGGCCCGGCTGTAGATCAGGTTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAA 2199
Qy    1524 CCCTGATCTNTCAG-CCCCAAAACACCCAATTGGCCCAAATTGGCCCACTGTNTGCCCCTGGGG 1582
           ||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    2200 CCCTGATCTCTCAGCCCCCAAAACACCTACCCAATTGGCCCAAATTGGCCCACTGTCTCTCCCCCGGGG 2259
Qy    1583 GTGTCAGAAATTNTAGCGCAAGGCACTTTAAAGCGTGGATTGTTTAAAGAGCTCTCCCAG 1642
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2260 GTGTCAGAAATTTCAGCGCAAGGCACTTTAAAAACGTGGATTGTTTAAAGAGCTCTCCAG 2319
Qy    1643 GCCCCACCAAGAGGGGTGGATCACACCTCAGTGGGAAGAAAAATAAAATTTTCCTTCAGGTT 1702
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2320 GCCCCACCAAGAGGGGTGGATCACACCTCAGTGGGAAGAAAAATAAAATTTTCCTTCAGGTT 2379
Qy    1703 TT 1704
           ||
Db    2380 TT 2381
```

RESULT 9
```
BD127811
LOCUS       BD127811          2780 bp   DNA    linear    PAT 18-SEP-2002
DEFINITION  Primer for synthesizing full-length cDNA and use thereof.
ACCESSION   BD127811
VERSION     BD127811.1   GI:23222756
KEYWORDS    JP 2002017375-A/3242.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2780)
  AUTHORS   Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
            Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
            Koga,H.
  TITLE     Primer for synthesizing full-length cDNA and use thereof
  JOURNAL   Patent: JP 2002017375-A 3242 22-JAN-2002;
```

REFERENCE   2 (bases 1 to 2780)
AUTHORS   Isogai,T. and Otsuki,T.
TITLE   Direct Submission
JOURNAL   Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
            Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3375, Fax:81-438-52-3986)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology; cDNA library construction:
            Institute of Medical Science, University of Tokyo, Laboratory of
            Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
            sequencing and clone selection: Helix Research Institute (supported
            by Japan Key Technology Center etc.).
FEATURES   Location/Qualifiers
     source   1..2780
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="NT2RP3000789"
                     /cell_line="NT2"
                     /cell_type="teratocarcinoma"
                     /clone_lib="NT2RP3"
                     /note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
                     precursor cells after 2-weeks retinoic acid (RA)
                     induction"
ORIGIN

Query Match   81.3%;   Score 1389;   DB 9;   Length 2780;
Best Local Similarity   99.2%;   Pred. No. 0;
Matches 1405;   Conservative 0;   Mismatches 5;   Indels 6;   Gaps 1;

```
Qy   299 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACGAGTCCAGCTGAA 358
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACGAGTCCAGCTGAA 60

Qy   359 AAAGCCATCAGTGTGCACTCCACCCTGAGGGTGCTGCTCCTCCGCTTGTAAGATGATCTTG 418
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AAAGCCATCAGTGTGCACTCCACCCTGAGGGTGCTGCTCCTCCGCTTGTAAGATGATCTTG 120

Qy   419 GAGATTATGCATAAAGAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTGAGATC 478
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAGATTATGCATAAAGAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTGAGATC 180

Qy   479 CTGGCCCATAATAACTTTGTAGGGCGTCCATTGGCAAGAGGAGGACGGAACTGAGAAG 538
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGGCCCATAATAACTTTGTAGGGCGTCCATTGGCAAGAGGAGGACGGAACTGAGAAG 240

Qy   539 GTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTTTAC 598
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTTTAC 300

Qy   599 AACCCTGAGGAGGACCATCGAAGGGGGGCCATCGAGAATTGTTGCAGGGCGAGCAG 658
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AACCCTGAGGAGGACCATCGAAGGGGGGCCATCGAGAATTGTTGCAGGGCGAGCAG 360

Qy   659 GAAATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGC----- 712
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GAAATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGCCTGCAG 420

Qy   713 TCTCACCTGATCCTGGCCTGAACCTGGCTGCTGTGTAGTGTCTTTTCCAGCTTCATCCAGC 772
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCTCACCTGATCCTGGCCTGAACCTGGCTGCTGTGTAGTGTCTTTTCCAGCTTCATCCAGC 480

Qy   773 GCAGTCCGCCGCCGCCCCAGCAGCGGTTACTGGGGCTGCTGGGGCGGCCGAGCAG 832
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GCAGTCCGCCGCCGCCCCAGCAGCGGTTACTGGGGCTGCTGGGGCGGCCGAGCAG 540

Qy   833 GCTCCCCGGAGCAGGAGATGGTGCAGGTGTTTATCCCCGGCCAGTGGGGCCCATCATC 892
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCTCCCCGGAGCAGGAGATGGTGCAGGTGTTTATCCCCGGCCAGTGGGGCCCATCATC 600

Qy   893 GGCAAGAAGGGGCAGCACATCAAAACAGCTCTCCCGGTTGCCAGCGCCTCATCAAGATT 952
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   661 GCACCACCCGAAACACCTGACTCCAAAGTTCGTATGATGGTTATCATCACTGGACCGCCAGAG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1013 GCCCAATTCAAGGCTCAGGGAGAATCTATGCAAACTCAAGGAGGAGAACTTCTTTGGT 1072
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GCCCAATTCAAGGCTCAGGGAGAAGAATCTATGCAAACTCAAGGAGGAGAACTTCTTTGGT 780

Qy  1073 CCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCCGG 1132
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCCGG 840

Qy  1133 GTCATTGGCAAAGGTGGAAAAACGGTGAACGAGTGCAGAATTGACGGCAGCTGAGGTG 1192
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GTCATTGGCAAAGGTGGAAAAACGGTGAACGAGTGCAGAATTGACGGCAGCTGAGGTG 900

Qy  1193 GTAGTACCAAGAGACCAGCACCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGA 1252
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GTAGTACCAAGAGACCAGCACCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGA 960

Qy  1253 CATTTCTATGCCAGTCAGATGGCTCAACGGTCAACCGAGAGATCCGAGACATCCTGGCCCAGGTTAAG 1312
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CATTTCTATGCCAGTCAGATGGCTCAACGGTCAACCGAGAGATCCGAGACATCCTGGCCCAGGTTAAG 1020

Qy  1313 CAGCAGCATCAGAAGGGACAGAGAGTAACCAGGCCCAGGCACGAGGAAGTGACCAGCCCCT 1372
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CAGCAGCATCAGAAGGGACAGAGAGTAACCAGGCCCAGGCACGAGGAAGTGACCAGCCCCT 1080

Qy  1373 CCCTGTCCCTTNGAGTCCAGGACACAACACACGGGCAGAAATCGAGAGTGTCCTCCCCGGC 1432
         |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CCCTGTCCCTTCGAGTCCGAGGACACAACACACGGGCAGAAATCGAGAGTGTCCTCCCCGGC 1140

Qy  1433 AGGGCCTGAGAAGTGAGAATGTTCGGGAATCCGGAACNTGGCCGGCTGTAGATCNAGTTTGCCC 1492
         |||||||||||||||||||||||||||||||||||| |||||||||||||||||| |||||||||
Db  1141 AGGGCCTGAGAAGTGAGAATGTTCGGGAATCCGGAACACCTGGGCCGGCTGTAGATCNAGTTTGCCC 1200

Qy  1493 ACTTGATTGAGAAAGATGTTCCAGTGAGGGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCACC 1552
         ||||||||||||||| |||||||||||||||||||||||||||||| |||||||||||||||||||
Db  1201 ACTTGATTGAGAAGAAGATGTTCCAGTGAGGGTGAGGAACCCTGATCTCTCAGCCCCAAACACCCACC 1260

Qy  1553 CAATTGGCCCAACACTGNTNGCCCCTCGGGGTGTCAGAAATTNTAGGCAAGGCACTTTT 1612
         ||||||||||| ||| | | |||||||||||||||| |||||| ||||||||||||||||
Db  1261 CAATTGGCCCCACCGTCTGCCCCTCGGGGTGTCATTGGCAAGGCACTTTT 1320

Qy  1613 AAACGTGGATTGTTAAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGATCACCACCTCAG 1672
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 AAACGTGGATTGTTAAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGATCACCACCTCAG 1380

Qy  1673 TGGGAAGAAAATAAAATTTCCTTCAGTTTTAAAA 1708
         |||||||| |||||||||||||||||||||||||
Db  1381 TGGGAAGAAGAAAATAAAATTTCCTTCAGTTTTAAAA 1416
```

RESULT 10
AK074915
LOCUS   AK074915   2780 bp   mRNA   linear   PRI 03-SEP-2002
DEFINITION Homo sapiens cDNA FLJ90434 fis, clone NT2RP3000789, highly similar
            to Homo sapiens IGF-II mRNA-binding protein 1 (IMP-1) mRNA.
ACCESSION   AK074915
VERSION   AK074915.1 GI:22760672
KEYWORDS   oligo capping; fis (full insert sequence).
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS   Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
            Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
            Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
            Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
            Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
            Ninomiya,K.
TITLE   NEDO human cDNA sequencing project
JOURNAL   Unpublished

```
                                                                /mol_type="unassigned DNA"

ORIGIN

Query Match          68.2%;    Score 1165;  DB 6;   Length 2224;
Best Local Similarity  82.8%;  Pred. No. 1.8e-282;
Matches 1411;  Conservative   0;  Mismatches 260;  Indels  33;  Gaps   6;

Qy    30 CCCGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT    89
            ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   525 CCAGGCAAGCTATCATGAGGCTAAATGGCCATCAACTGGAGAACCATGCCCTGAAGGTCT   584

Qy    90 CCTACATCCCGATGAGCAGATAGCACAGGGACCTGAGGATGGGCGGCGGGGCCCAAGCAGC   149
            ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   585 CCTACATACCTGATGAGCAGATAACACAAGGTCCTGAGGATGGGCGTCGTGGAGGCTTTG   644

Qy   150 GCTCTCGGGGTCAGCCCGCCAGGGCTCACCTGTGCAGCGGGCTCCAGCCAAGCAGC   209
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   645 GGTCTCGGGGCCAGCCCGGCCAGGGCTCGCCCGTGCAGCGGGCTCCAGCCAAGCAGC   704

Qy   210 AGCAAGTGGACATCCCCTCGGCTCCTGGTGTGGCCACCCAGTGGTGGTGCCATTATTG   269
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   705 AGCCAGTGGACATCCCCTCCTGGTGCCTCCTGGTGTGGCCACCCAGTGGTGGTGCCATCATTG   764

Qy   270 GCAAGGAGGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGC   329
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   765 GCAAGGAGGGGTGCCACCATCCGCAAGATCACAAAACATCACAAAATAGACGTGC   824

Qy   330 ATAGGAAGGAGAAGCCAGGTGCAGCTGAAGCCATCAGTGTGCACTCCACCCCTGAGG   389
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   825 ATAGGAAGGAGAATGCGGGCGCCATCAGCGTGCATTCAACCCCTGAAG   884

Qy   390 GCTGCTCCTCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA   449
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   885 GCTGCTCCTCGCGTGCAAGATGATCTTGGAGATTATGCACAAGGAGGCAAAGGACACCA   944

Qy   450 AAACGGCTGACGAGTTCCCCTGAAGATCCTGGCCCCATAATAACTTTGTAGGGCGTCTCA   509
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   945 AAACGGCAGATGAAGTTCCCCTGAAGATCCTGGCTCATAACAACTTCGTGGGGACTCA   1004

Qy   510 TTGGGCAAGGAAGGACGAACCTTACACCTGAGAGGACCATCACTGTGAAGGGG   569
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1005 TTGGGCAAGGAAGGACGAACCTGAGAGGAGGACGAAGATCACCA   1064

Qy   570 TCTCCGTTGCAAGACCTTACCCTTCAGAAATTNTAGGCCAGGCCACTTTT   629
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1065 TCTCATCGCTCCAGACCTTCAGCCCAAGGTGATCGAAGGGG   1124

Qy   630 CCATCGAGAATTGTTGCAGGGCCGAGCAGCAGAAATAATGAAGAAGTTCGGGAGGCCTATG   689
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1125 CCATTGAGAACTGTTGCAGGGCCATGAGGAGATCATGAAGAAGTTCGAGAGGCCTTACG   1184

Qy   690 AGAATGATGTGGCTGCCATGAGC------TCTCACGTGATCCCTGGCCTGGCTG   743
            ||||||||||||||||||||||||      |||||||||||||||||||||||||
Db  1185 AGAACGACGTGGCCGCCATGAGCTTGCAGTGCCATCCCTGGGCCTTAACCCGGCTG   1244

Qy   744 CTGTAGGTCTTTCCCAGTTCATCCCAGCTT   803
            |||||||||||||||||||||||||||
Db  1245 CTGTAGGTCTCTTCCCAGCTTCATCCCAGCTTCCTCCTCCCGCCGCCAGCGTTACTG   1304

Qy   804 GGGCTGCTCCCTATAGTCTCTTCATGCAGGCCTCGCAGGCTCCGGGAGAGGATGGTGCAGTGTTA   863
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1305 GGGCTGCTCCCTATAGTCTCTTCATGCGAGGTCCGGGACCCAGGAGATGGTACAAGTGTTCA   1364

Qy   864 TCCCCCCAGGCAGTGGGCACCATCATCCGCAAGAAGGGGCAGCACATCAAACAGCTCT   923
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1365 TCCCCCCAGGCAGTCGTGGGCCATCATTGGCAAGAAGAGGGCCAGCACATCAAACAAACTCT   1424

Qy   924 CCCGGTTGCCACGGCCTCATCAAGATTGCACCACCTGAAGATGTCACAACCTGACTCCAAAGTTC   983
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1425 CCCGTTTCGCCACGGCCTCCATCAAGATTGCTCCACCCAAGAAAACACCTGACTCCAAAGTTC   1484

Qy   984 GTATGGTTATCATCACTGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATG   1043
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGTTTGCAGCGCCTCCATCAAGATT   660

Qy   953 GCACCACCCGAAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCGCCAGAG   1012
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GCACCACCCGAAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCGCCAGAG   720

Qy  1013 GCCCAATTCAAGGCTCAGGGAGCAAGGAATTCAGGAGGAGAGAACTTCTTTGGT   1072
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GCCCAATTCAAGGCTCAGGGAAGAATTCAGGAGGAGAGAACTTCTTTGGT   780

Qy  1073 CCCAAGGAGGAAGTGAAGCTGGAAGCGGAGACCCACATACGTGTGCCAGTCAGCTGCCGCGG   1132
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CCCAAGGAGGAAGTGAAGCTGGAGAATCATGGCAAGCTCAAGAGCATCGTGCCAGTCGGCCGCGG   840

Qy  1133 GTCATTGGCAAAAGTGGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTG   1192
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GTCATTGGCAAAAGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACGCGCAGCTGAGGTG   900

Qy  1193 GTAGTACCAAGAGAGACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGA   1252
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GTAGTACCAAGAGAGACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGA   960

Qy  1253 CATTTCTATGCCAGTCAGATGGCTCAACGGAGACATCCGAGACATCCTGGCCCCAGGTTAAG   1312
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CATTTCTATGCCAGTCAGATGGCTCAACGGAGGACATCCTGGCCCCAGGTTAAG   1020

Qy  1313 CAGCAGCATCAGAAGGACAGAGTAACCAGCGCCAGGGAAGTGACCAGCCCCT   1372
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CAGCAGCATCAGAAGGACAGCAGGCCAGGGAAGTGACCAGCCCCT   1080

Qy  1373 CCCTGTCCTTNGAGTCAGGACAACGGGCAGAAATCGAGAGTGGCTCTCCCCGGC   1432
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 CCCTGTCCTTCGAGTCAGGACAACGGGCAGAAATCGAGAGTGGCTCTCCCCGGC   1140

Qy  1433 AGGCCTGAGAATGAGTGGGAATCCGGGACACNTGGGCGGCGGTTGCCC   1492
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 AGGCCTGAGAATGAGTGGGAATCCGGGACACCTGGGCGTGTAGACTGGTTTGCCC   1200

Qy  1493 ACTTGATTGAGAAGAGATGTTCCAGTGAGGAGACCCTGATCTNTCAGCCCCAAACACCCACC   1552
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 ACTTGATTGAGAAGAGATGTTCCAGTGAGGAGACCCTGATCTCTCAGCCCCAAACACCCACC   1260

Qy  1553 CAATTGGCCCAACACTGTNTGCCCTCGGGGTGTCAGAAATTNTAGGCGCAAGGCACTTTT   1612
            ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  1261 CAATTGGCCCACACTGTCTGCCCTCGGGGTGTCAGAAATTCTAGGCGCAAGGCACTTTT   1320

Qy  1613 AAACGTGGATTGTTTAAAGAGCTCTTCCAGGGCCCACCAAGAGGGTGGATCACACCTCAG   1672
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 AAACGTGGATTGTTTAAAGAGCTCTTCCAGGCCCCACCAAGAGGGGTGGATCACACCTCAG   1380

Qy  1673 TGGGAAGAAAATAAAATTTCCTTCAGGTTTTAAAA   1708
            |||||||||||||||||||||||||||||||||||
Db  1381 TGGGAAGAAAATAAAATTTCCTTCAGGTTTTAAAA   1416

RESULT 11
AR160244
LOCUS       AR160244               2224 bp    DNA       linear       PAT 17-OCT-2001
DEFINITION  Sequence 1 from patent US 6255055.
ACCESSION   AR160244
VERSION     AR160244.1  GI:16223887
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2224)
  AUTHORS   Ross,J.
  TITLE     c-myc coding region determinant-binding protein (CRD-BP) and its
            nucleic acid sequence
  JOURNAL   Patent: US 6255055-A 1 03-JUL-2001;
FEATURES             Location/Qualifiers
     source          1..2224
                     /organism="unknown"
```

```
Db   1485  GAATGGTCGTCATCACTGGACCCCCAGAGGCTCAGTTCAGGCTTCAAGGCTGAGGAAGAATTTATG  1544
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1044  GCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGGTGGAGACCCCACA  1103
                 ||||  |||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1545  GCAAACTAAAGAGAGAGAATTTCTTTGGTCCCAAGGAGGAAGTGAAAGGTAGAGACCCACA  1604

Qy   1104  TACGTGTGCCAGCATCAGCAGCTCGCCGCGGTCATTGGCAAAGGTGGAAAAACGGTGAACG  1163
                 |||  |||||||||||||||||| || |||  |||||||| || |||||||||||||||||
Db   1605  TACGGGTTCCGGCTTCAGCGAGCGCCGGCCGTCATCGGCACAAACGGCGCAAAACGGTGAATG  1664

Qy   1164  AGTTGCAGAATTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGA  1223
                 ||||||||||||||| ||||| |||||  |||||||||||||||||||||||||||||||
Db   1665  AGCTGCAGAACTTGACTGCAGCTGAGGTGGTAGTGCCAAGAGACCAGACCCCGATGAGA  1724

Qy   1224  ACGACCAGGTCATCGTGAAAATCTGGACATTTCTATGCCAGTCAGATGGCTCAACGA  1283
                 |||||||||||||||||  ||   || |||  |||||||||||||||||||| |||||||
Db   1725  ACGACCAAGTCATTGTTAAGATCAATCGGACATTTCTATGCCAGATGGCTCAGCGA  1784

Qy   1284  AGATCCGAGACATCCTGCCCAGTTAAGCAGCATCAGAAGGGACAGAGTAACCAGG  1343
                 |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1785  AGATCCGAGACATCCTGGCTCAAGTAAGCAACAGCACCAGAGGGACAGAGCAACCTGG  1844

Qy   1344  CCCAGGCACCGAGGAAGTGA-CCAGCCCCTCCCTTGTCCCTTNGAGTCCAGGACAACACG  1402
                 ||||||||||||||||||||| |||||||||||||||||||  |||||||||||||||||||
Db   1845  CCCAGGCACCGAGGAAGTGACCCGCCCGTCCCATTGGCTCCAGATCAGCAGG  1904

Qy   1403  GGCAGAA------------ATCGAGAGTGTGCTCTCCCCGGCAGGCCTCAG  1441
                 |||||||            |||||||||||||| |||| |||||||||||||
Db   1905  AGGAACACAGAACTGGAGGGGGCGGTGTGGGGGGCGGTGTGTGTTTTCCCAGCAGGCCCTGAG  1964

Qy   1442  AATGAGTGGGAATCGGGACACNTGGGCGCGGGTGTAGATACAGGTTTGCCCACTTGATTG  1501
                 |||||||||||||||| | || ||||||||||| ||||||||||||||||||||||||||
Db   1965  AATGAGTGGGAATCAGGG-CATTTGGGCGCGGGTGGCTGGAGATCAGGTTTGCACACTGTATTG  2023

Qy   1502  AGAAAGATGGTCCAGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCAATTGGCC  1561
                 |||||| |||||||| |||||||||||||| |||||||||||||  ||||||||||||
Db   2024  AGACAATGTTCCAGTGAGGAATCCTGATCTCTCGCCCCCAA--TTGAGCCAGCTGCCA  2081

Qy   1562  CAACACTGTNTGCCCCTCGGGGTGTCAGAAATNTAGCCCAAGGCACTTTTAAACGTGA  1621
                 ||||| |||| ||||| |||||||   |||| ||||  |||||||||| ||||||||
Db   2082  CAGCCCACCCCTGAAATATCACCATTGCAATCATAGCTTGGGTTGCTTTTAAACGTGA  2141

Qy   1622  TTGTTTAAAGAAGTCTCCAGGCCCACCAAAGAGGGTGATCACACCTCAGTGGGAGAA  1681
                 |||||  | |||||||||||| |||||||||||||||||||||||||||||||||||||
Db   2142  TTGTCT--TGAAGTTCTCCAGCCTCCATGCGGAAGGATGGGTCAGATCCCAGTGGGGAGAGAG  2199

Qy   1682  AAATAAAATTTCCTTCAGTTTTA  1705
                 ||||||||||||||||||||||||
Db   2200  AAATAAAATTTCCTTCAGTTTTA  2223
```

```
RESULT 12
LOCUS       AF061569              2223 bp    mRNA     linear     ROD 11-NOV-1998
DEFINITION  Mus musculus coding region determinant binding protein mRNA,
            complete cds.
ACCESSION   AF061569
VERSION     AF061569.1  GI:3273748
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2223)
  AUTHORS   Bernstein,P.L., Herrick,D.J., Prokipcak,R.D. and Ross,J.
  TITLE     Control of c-myc mRNA half-life in vitro by a protein capable of
            binding to a coding region stability determinant
  JOURNAL   Genes Dev. 6 (4), 642-654 (1992)
  MEDLINE   92217743
  PUBMED    1559612
REFERENCE   2  (bases 1 to 2223)
  AUTHORS   Herrick,D.J. and Ross,J.
  TITLE     The half-life of c-myc mRNA in growing and serum-stimulated cells:
```

```
            influence of the coding and 3' untranslated regions and role of
            ribosome translocation
  JOURNAL   Mol. Cell. Biol. 14 (3), 2119-2128 (1994)
  MEDLINE   94158886
  PUBMED    8114742
REFERENCE   3  (bases 1 to 2223)
  AUTHORS   Prokipcak,R.D., Herrick,D.J. and Ross,J.
  TITLE     Purification and properties of a protein that binds to the
            C-terminal coding region of human c-myc mRNA
  JOURNAL   J. Biol. Chem. 269 (12), 9261-9269 (1994)
  MEDLINE   94179348
  PUBMED    8132663
REFERENCE   4  (bases 1 to 2223)
  AUTHORS   Leeds,P., Kren,B.T., Boylan,J.M., Betz,N.A., Steer,C.J.,
            Gruppuso,P.A. and Ross,J.
  TITLE     Developmental regulation of CRD-BP, an RNA-binding protein that
            stabilizes c-myc mRNA in vitro
  JOURNAL   Oncogene 14 (11), 1279-1286 (1997)
  MEDLINE   97322234
  PUBMED    9178888
REFERENCE   5  (bases 1 to 2223)
  AUTHORS   Doyle,G.A., Betz,N.A., Leeds,P.F., Fleisig,A.J., Prokipcak,R.D. and
            Ross,J.
  TITLE     The c-myc coding region determinant-binding protein: a member of a
            family of KH domain RNA-binding proteins
  JOURNAL   Nucleic Acids Res. 26 (22), 5036-5044 (1998)
  MEDLINE   99035044
  PUBMED    9801297
REFERENCE   6  (bases 1 to 2223)
  AUTHORS   Ross,J., Prokipcak,R.D., Leeds,P., Doyle,G.A.R., Betz,N.A. and
            Fleisig,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-1998) McArdle Laboratory, University of Wisconsin
            - Madison, 1400 University Ave., Madison, WI 53706, USA
FEATURES             Location/Qualifiers
     source          1..2223
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     CDS             131..1864
                     /function="putative c-myc mRNA binding protein"
                     /note="CRD-BP"
                     /codon_start=1
                     /product="coding region determinant binding protein"
                     /protein_id="AAC72743.1"
                     /db_xref="GI:3273749"
                     /translation="MNKLYIGNLNESVTPADLEKVFAEHKISYSGQFLVKSGYAFVDC
                     PDEHWAMKAIETFSGKVELQGKRLEIEHSVPKKQRSRKIQRNIPPQLRWEVLDSLLA
                     QYGTVENCEQVNTESETAVVNVTYSNREQTRQAIMKLNGHQLENHALKVSYIPDEQIT
                     QGPENGRRGGFGSRGCPRQSPVAAGAPAKQQPVDIPLRLLVPTQYGVAIIGKEGATI
                     RNITKQTQSKIDVHRKENAGAAEKAISVHSTPEGCSSACKMILEIMHKEAKDTKTADE
                     VPLKILAHNNFVGRLIGKEGRNLKKVEQDTETKITISSLQDLTLYNPERITTVKGAIE
                     NCCRAEQEIMKKVREAYENDVAAMSLQSHLIPGLNLAAVGLFPASSSAVPPPSSVTG
                     VRMVVITGPPEAQFKAQGRIYGKLKEENFGPKEKVKLETHIRVPASAAGRVIGKKGK
                     TVNELQNLTAAEVVVPRDQTPDENDQVIVKIIGHFYASQWAQRKIRDILAQVKQQHQK
                     GQSNLAQARRK"
```

```
ORIGIN
Query Match             68.1%;  Score 1162.6;  DB 10;  Length 2223;
Best Local Similarity   83.1%;  Pred. No. 7.2e-282;
Matches 1416;  Conservative    0;  Mismatches 254;  Indels 34;  Gaps    7;

Qy      30  CCCGGGGAGCCAGCCATCATCAGAAGTGAATGGCCACCAGTTGGAGAACCCTGGAGAACGTCT   89
                ||||  |||||||||||||||||||||||||||||||  || |||||||||||||||||||
Db     525  CCAGGCAAGCTATCATCAGAAGTCAAATGGCCATCAACTGGAGAACCATGCCCTGAAGGTCT  584

Qy      90  CCTACATCCCCGATGAGCAGATAGCACAGGGACCTGAGAATGGGCCGCCGAGGGGGCTTTG  149
                ||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     585  CCTACATACCTGATGAGCAGATAACACAAGGTCCTGAAGGTCCTGAGAATGGGCCTTTG  644

Qy     150  GTCTCCGGGTGCAGCCCCCGCCAGGGCTCACCTGTGCAGCCAAGCAGC  209
                ||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy    690 AGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGACCTGCTG 743
              ||||||||||||||||||||||      ||||||||||||||||||||||||
Db   1185 AGAACGACGTGGGCCGCCATGAGCTTGCAGTGAGCTTCATCCCTGGGCTTAACCTGGCTG 1244

Qy    744 CTGTAGGTGTCTTTCCCAGCTTCATCCAGCGCGCTGTCCCTGCGCTGTCACTG 803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1245 CTGTAGGTGTCTTTCCCAGCTTCATCCAGCGCGCTGTCCCTGCGCTGTCACTG 1304

Qy    804 GGGCTGCTCCCTATAGCTCCTTTATGCAGCAGGAGATGGTGCAGGTGTGTTA 863
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1305 GGGCTGCTCCCTATAGCTCCTTCCTCATGCAGGAGGAGATGGTGCAAGTGTTCA 1364

Qy    864 TCCCCGCCCAGTGGGCCGCCATCGGCGACCACCATCAAACAGCTCT 923
              ||||||||||||||||||||||||||||||||||||||||||||||
Db   1365 TCCCCGCCCAGGCGTGGGCGCCCATCATTGGCGACCACCATCAAACAACTCT 1424

Qy    924 CCCGGTTGCCAGCGCCTTCCAATCAAGATTGCACCACCTGACTCCAAAGTTC 983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    983 CCCGTTTGCCAGCGCCTTCCATCCAGATTGCTCCACCAACCTGACTCCAAAGTTC 1484

Qy    984 GTATGGTTATCATCACTGACACGCCCAGAGCCCAATTCAAGGCTCAGGAGAATCATG 1043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1485 GAATGGTCGTCATCACTGACACTGACCGTCAGTCAAGGCTCAGGAGAATTTATG 1544

Qy   1044 GCAAACTCAAGGAGGAGGAGAACTTCTTTGTCCCAAGAGGAGGAGGAGACCCACA 1103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1545 GCAAACTAAAGAAGAGAATTTCTTTGTCCCAAGGAGGAAGCTAGAGACCCACA 1604

Qy   1104 TACGGTGCCAGCATCAGCAGCTGGCCGGCCGGCCGCCGTCATTGGCATCGGCAAAAGGAACG 1163
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1605 TACGGGTTCCGGCTTCAGCAGTTCAGCAGCTCGGCCGCCGCCGGCCGCCAAAAGCGGAATG 1664

Qy   1164 AGTTGCAGAATTGACGCGGCATCGTGAAAATCATCGACAGTGGAGTGGTAGTGCCAGTGATGAGA 1223
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1665 AGCTGCAGAACTGACTGCGTTCAGCGAGTGGTAGTGCCAGTGGTAGTGCCAGTGATGAGA 1724

Qy   1224 ACGACCAGTCATCGTGAAAATCATCTATGCCAGTCAAGGAGGAACACCCCTGATCGAGA 1283
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1725 ACGACCAGTCATCGTGAAGATCATCGGACATTCTATGCCAGCCAGATGGCTCAGCGGA 1784
```

```
Db    645 GGTCTCGGGGGCAGCCCCGGCAAGGGTCGCCCGTGGCGAGCAGCAAGCAGC 704
Qy    210 AGCAAGTGACATCCCCCTTCGGCTCGGTGTGGTGCATTATTG 269
              ||||||||||||||||||||||||||||||||||||||||||
Db    705 AGCCAGTGACATCCCTCTCCGGCTCTCCGGTGTGCCAGTGTATGTGTATCATTG 764

Qy    270 GCAAGGAGGGGCCACCATCACAAATCACAAAACAGACCCAGTCCAAGATAGACGTGC 329
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    765 GCAAGGAGGGTGCCACCATCACAAATCACAAAACAGACCAGTCCAAAATAGACGTGC 824

Qy    330 ATAGGAAGGAGACGCAGGTGCGAGCTGAAAAAGCCATCAGTGCACTCCACCCCTGAGG 389
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    825 ATAGGAAGGAGAATGCGGGCGTGCGAGGACGGCCATCAGCCGTGCATTCAACCCCTGAAG 884

Qy    390 GCTGCTCCTCGCTTGTAAGATGATCTTGGAGATTATTGCATAAAGAGGCTAAAGGACACCA 449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    885 GCTGCTCCTCCCGCGTGCAAGAAGATGATCTTGGAGATTATTGCACAAGGAGGCAAAGGACACCA 944

Qy    450 AAACGGCTGACGAAGTTCCCTGAAGATCCTGCCCATAATAACTTGTTAGGGCGTCTCA 509
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    945 AAACGGCAGCAGTGAAGTCCCTGAAGATCCTGCTCATAACAACTTCGTCGGGCGACTCA 1004

Qy    510 TTGGCAAGGAAGGACGGAACCTGTAGAGACGGAAGGTAGAGACAAGAAGACCGAGACAAAATCACCA 569
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1005 TTGGCAAGGAAGGCCGGAACCTGGAAGAAGGTGAGCAGCAGCAACAGAGATCACCA 1064

Qy    570 TCTCCTCGTTGCAAGAACTTACCCTTTACAACCCTGAGGAGGACCATCACTGTGAAGGGGG 629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1065 TCTCATCGCTCCAGGAACCTCAGCTCTATAAACCCTGAGGAGGACCATCACTGTGAAGGGCG 1124

Qy    630 CCATCGAGAATTGTTGCAGCTGCCATGAGTGTTCGGGGAGGTCGGGGAGGCTATG 689
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1125 CCATTGAGAACGTGTTGCAGCTGCCATGAGTGTTCGGGAGAGGAGAAGTTCGGAGAGGCTTACG 1184
```

```
Db   1284 AGATCCGAGACATCTCTGGCCTTAAGCAGCAGCAGCATCAGAAGGAGGACGAGAGTAACCAGG 1343
Db   1785 AGATCCGAGACATCTCTGGCTCAAGTTAAGCAACAGCAGCACCAGCAGGAGCAGCAACCTGG 1844

Qy   1344 CCAGGCACGGAGGAAGTGA-CCAGCCCTCCTCCTTGAGTCCAGGACACAACG 1402
              ||||||||||||||||||  |||||||||||| ||||||||||||||||||
Db   1845 CCCAGGCTCGGAGGAAGTGACCCCGGCCCCTCCTCGTCCAAGATCAGCAGG 1904

Qy   1403 GGCAGAA------------------ATCGAGAGTGTGCTCTCCCCGGCAGGCCTGAG 1441
              |||||||                  |||||||||||||||||||||||||||||||
Db   1905 AGGAACACAGAACTCGGAGGGGCGGGGGCGGGGTGTGTGTTTTTCCCAGCAGGCCTGAG 1964

Qy   1442 AATGAGTGGGAATCCGGGACNTGGGGCCGGGGCGTCTGAGATCAGGGTTGCCACTTGATTG 1501
              ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   1965 AATGAGTGGGAATCAGGG-CATTTGGGCCTGGCTGGAGATCAGGTTTGCACACTGTATTG 2023

Qy   1502 AGAAAGATGTTCCAGTGAGGAACCCTGATCNTNCAGCCCCAAACACCCACCCATTGGCC 1561
              ||||||||||||||||||||||||||||| | ||||||||||   |||||||||||||
Db   2024 AGAACATGTTCCAGTGAGGAGGATCCTGATCCTCTCGCCCCCCAA--TTGAGCCAGTGGCCA 2081

Qy   1562 CAACACTGTNTGCCCCTCGGGTGTCAGAAATTNTAGCGCAAGGCACTTTTAAACGTGGA 1621
              |||||||| |||||||||||||| ||||||| |||||||||||||||||||||||||||
Db   2082 CAGCCCACCCTTGGAATATCCAATCATTGCAATCATAGCTTGGGTTGCTTTTAAACGTGGA 2141

Qy   1622 TTGTTAAAGAGCTCCAGCCCCACCAAGAGGGTGGATCACACCTCAGTGGGAAGAA 1681
              ||||| |||||||||||||||||||||||||||||||||  |||||||||||||
Db   2142 TTGTCT--TGAAGTTCTCCAGCCTCCATGGA-AGGATGGGTCAGATCCCAGTGTGGGGAAGA 2198

Qy   1682 AAATAAAATTCCTTCAGTTTTTA 1705
              ||||||||||||||||||||||
Db   2199 GAATAAAATTTCCTTCAGTTTTTA 2222
```

RESULT 13
BC051679
LOCUS   BC051679   2444 bp   mRNA   linear   ROD 08-OCT-2003
DEFINITION Mus musculus insulin-like growth factor 2, binding protein 1, mRNA
            (cDNA clone MGC:60613 IMAGE:30008106), complete cds.
ACCESSION   BC051679
VERSION   BC051679.1 GI:30354043
KEYWORDS   MGC.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 2444)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE   Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   MEDLINE 22388257
   PUBMED   12477932
REFERENCE   2 (bases 1 to 2444)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission

JOURNAL   Submitted (30-APR-2003) National Institutes of Health, Mammalian
           Gene Collection (MGC), Cancer Genomics Office, National Cancer
           Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
           USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Dr. Kenneth R. Boheler (National Insitute on
           Aging, USA)
           cDNA Library Preparation: Yulan Piao and Minoru Ko (National
           Institute on Aging, NIH: http://lgsun.grc.nia.nih.gov/cDNA/)
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Baylor College of Medicine Human Genome
           Sequencing Center
           Center code: BCM-HGSC
           Web site: http://www.hgsc.bcm.tmc.edu/cdna/
           Contact: amg@bcm.tmc.edu
           Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
           Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
           A.N., Gibbs, R.A.

           Clone distribution: MGC clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Series: IRAK Plate: 111 Row: e Column: 24
           This clone was selected for full length sequencing because it
           passed the following selection criteria: matched mRNA gi: 6753517.
FEATURES   Location/Qualifiers
     source   1..2444
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="129/Sv x 129/Sv-CP"
                     /db_xref="taxon:10090"
                     /clone="MGC:60613 IMAGE:30008106"
                     /tissue_type="ES Cell, undifferentiated, mouse, R1
                     Embryonic stem cells"
                     /clone_lib="NIA Mouse Undifferentiated ES Cell cDNA
                     Library (Long)"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1"
     gene   1..2444
                     /gene="Igf2bp1"
                     /note="synonyms: IMP-1, Neilsen, CRD-BP"
                     /db_xref="LocusID:140486"
                     /db_xref="MGI:1890357"
     CDS   295..2028
                     /codon_start=1
                     /product="insulin-like growth factor 2, binding protein 1 "
                     /protein_id="AAH51679.1"
                     /db_xref="GI:30354044"
                     /db_xref="LocusID:140486"
                     /translation="MNKLYIGNLNESVTPADLEKVFAEHKISYSGQFLVKSGYAFVDC
                     PDEHWAMKAIETFSGKVELQGKRLEIEHSVPKKQRSRKIQRNIPPQLRWEVLDSLLA
                     QYGTVENCEQVNTESETAVVNVTYSNREQTRQAIMKLNGHQLENHALKVSYIPDEQIT
                     QGPENGRRGGFGSRGQPRQSGSPVAAGAPAKQQPVDIPLRLLVPTQYVGAIIGKEGATI
                     RNITKQTQSKIDVHRKENAGAAEKAISVHSTPEGCSSACKMILEIMHKEAKDTKTADG
                     VPLKIAHNNFVGRLIGKEGRNLKKVEQDTETKITISSLQDLTLYNPERTITVKGAIE
                     NCCRAEQEIMKKVREAYENDVAAMSLQSHIJPGLNLAAVGLFPASSSAVPPPSSVTG
                     AAPYSSFMQAPEQEMVQVFIPAQAVAIIGKKGQHIKQLSRFASASIKIAPPETPDSK
                     VRMVITGPPEAQFKAQGRIYGKLKEENFFGPKEEVKLETHIRVPASAAGRVIGKGGK
                     TVNELQNLTAAEVVVPRDQTPDENDQVIVKIIGHFYASQMAQRKIRDILAQVKQQHQK
                     GQSNLAQARRK"
     misc_feature   538..750
                     /note="RRM; Region: RNA recognition motif"
                     /db_xref="CDD:smart00360"
     misc_feature   877..1086
                     /note="KH; Region: K homology RNA-binding domain"
                     /db_xref="CDD:smart00322"
     misc_feature   1504..1719
                     /note="KH; Region: K homology RNA-binding domain"
                     /db_xref="CDD:smart00322"
ORIGIN

Query Match   68.0%; Score 1161.6; DB 10; Length 2444;

Best Local Similarity   82.6%;   Pred No. 1.3e-281;
Matches 1410; Conservative   0; Mismatches 264; Indels   33; Gaps   6;

Qy   30 CCCGGGGAGCCATCATGAAGCTGAATGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT   89
          ||||||| |||||||| ||||| || |||| ||||||||||||||||||||||||||
Db   689 CCAGGCAAGCTATCATGAGCTAAATGCCATCATCACTGGAGAACCATGCCCTGAAGGTCT 748

Qy   90 CCTACATCCCGATGAGCAGATAGCACAGGAGACCTGAGAATGGCGCCGAGAGGGGCTTTG 149
          |||||||||||||||||||| |||||||||||| |||||||| ||||||||||||||||
Db   749 CCTACATACCTGATGAGCAGATAACGCAAGGTCCTGAGAATGGCGTCGTGGAGGCTTTG 808

Qy   150 GCTCTCGGGTCAGCCCCGCCAGGGCTCACCTGTGCGAGCGGGGCCCAGCCAAGCAGC   209
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   809 GGTCTCGGGGCCAGCCGCCAAGGGCTCGCCCGTGCGAGCAGGGCTCCAGCCAAGCAGC 868

Qy   210 AGCAAGTGACATCCCCTTCGGCTCTTGTAGATGATCTTGGACATTATGCATAAAGAGGCTAAGGACACCA 269
          ||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||
Db   869 AGCCAGTGACATCCTCTCCGGGTCTCCTGTGGCTACGCAGTATGTAGGCGCTATCATTG 928

Qy   270 GCAAGGAGGGGCGCAACATCCGCAACATCACAAACAGACCCAGTCAAGATAGACGTGC   329
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   929 GCAAGGAGGGTGCCACCATCCGCAACATCCAAAACATCACAAACAGACGCAGTCAAGATAGACGTGC   988

Qy   330 ATAGGAAGGAACGCAGTGCAGTGGAGTTCCCCTGAAGAATCCTGGCTCATAACAACTTCGTCGGGCGACTCA 389
          ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   989 ATAGGAAGGAGAATGCGGGGCGTGCGGCGGAAGCCATCAGCGTGCATTCAACCCCTGAAG 1048

Qy   390 GCTGCTCTCCCGCTTGTGAAGATGATCTTGGACATTATGCATAAAGAGGCTAAGGACACCA 449
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1049 GCTGCTCTCCCGCTGCAAGATGATCTTGGAGATTATGCACAAGGAGGCAAAGGACACCA 1108

Qy   450 AAACGGCTGACAGAGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGCGTCTCA 509
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db   1109 AAACGGCAGATGAGTTCCCCTGAAGATCCTGGCTCATAACAACTTCGTCGGGCGACTCA 1168

Qy   510 TTGGCAAGGAGGGACGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAAATCACCA 569
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1169 TTGGCAAGGAGGGTCGGAACCTGAAGAAGGTGGACCAGGACACCAGAGCAATCACCA 1228

Qy   570 TCTCCTCGTTGCAAGACCTTACCCTTTACAACCCGAAGAGGACACTACTGTGAAGGGGGG 629
          |||||||| |||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1229 TCTCATCGCTCCAGGACCTTCACCCTCTATAACCCTGAGAGGACCATCACTGTGAAGGGCG 1288

Qy   630 CCATCGAGAATTGTTGCAGGGCCGAGCAGGAAATAATGAAGAAGTTCGGAGGGCCTATG 689
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||| |
Db   1289 CCATTGAGAACTGTTGCAGGGCCGAGCAGGAGATCATGAAGAAGTTCGAGAGGCTTACG 1348

Qy   690 AGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCCTGAACCTGGCTG 743
          ||||||| |||||| |||||||| |||| | ||||||||||| |||||||||||
Db   1349 AGAACGACGTGGCCGCCATGAGCCTGCAGTCCCATCCCTGGCCTTAACCTGGCTG 1408

Qy   744 CTGTAGGTCTTTTCCCAGCTTCATCGAGGCTCCCAGCAGTCCCGCCGCCGCCGCGGTTACTG 803
          |||||| ||||| |||||||||||||||||||| ||||||||||||| ||||||||||||
Db   1409 CTGTAGGTCTCCTCCCAGCTTCATCCAGCGTGTCCTCCTCCCCAGCAGTGTCACCG 1468

Qy   804 GGGCTGCTCCCTATAGCTCCTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTA 863
          |||||||||||| |||||| |||||| ||||||||||||||||||||||| ||||||
Db   1469 GGGCTGCTCCCTATAGCTCCTTCATCATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTCA 1528

Qy   864 TCCCCGCCAGCAGTGGCGCTGTGGCGCCATCATCGGCAAGAAGGGCAGCACATCAAACAGCTCT 923
          ||||||||||||| |||||| |||||||||||||||||||||||||||| |||||||||||
Db   1529 TCCCCGCCAGCAGCGGTGTGGCGCCATCATTGGCAAGAAGGGGCCAGCACATCAAACAACTCT 1588

Qy   924 CCCGGTTGCCAGCGCCTCCATCAAGATTGCACCACCGAAACACCTGACTCCAAAGTTC 983
          |||||||| ||||| |||||||||||||||||||||||||||||| ||||||||||||
Db   1589 CCCGGCTTGCCAGCGCCTCCATCAAGATTGCACCACCGAAACACCTGACTCCAAAGTTC 1648

Qy   984 GTATGGTTATCATCACTGGACCGCCGCCAATTCAGGGACCCCAATTCAAGGCTCAGGGAGAATCTATG 1043
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1649 GAATGGTCGTCATCACTGGACCCCCAGAGGCTCAGTTCAAGGCCCAGGGAGAATCTATG 1708

Qy   1044 GCAAACTCACGAGAGAGCACTTCTTTGGTCCAAGGAGGAAGTGAAGCTGGAGACCCACA 1103
          |||||||| |||||||||| || |||||||||||||||||||||||||||||||||||

```
FEATURES             Location/Qualifiers
     source          1..2105
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     CDS             223..1956
                     /note="rZBP1"
                     /codon_start=1
                     /product="b-actin zipcode binding protein 1"
                     /protein_id="AAO16210.1"
                     /db_xref="GI:27464838"
                     /translation="MNKLYIGNLNESVTPADLEKVFAEHKISYSGQFLVKSGYAFVDC
                     PDEHWAMKAIETFSGKVELQGKRLEIHSVRSKIQIRNIPPQLRWEVLDSLLA
                     QYGTVENCEQVNTESETAVVNVTYSNREQTRQAIMKLNGHQLENHALKVSYIPDEQIA
                     QGPENGRRGGFGSRGQPROQSPVAAGAPAKQQVDIPLRLLVPTQYVGAIIGKEGATI
                     RNITKQTQSKIDVHRKENAGAAEKAISVHSTPEGCSSACKMLEIMHKEAKDTKADE
                     VPLKILAHNFVGRLIGKEGRNLKKVEQDTEYKITISSLQDIJTLYNPERTITVKGAIE
                     NCCRAEOEIMKKVREAYENDVAAMSLQSHLIPGLNLAAVGLFPASSAVPPPSSVTG
                     AAPYGSFMQAPEQEMVQVFIPAQAVGAIIGKKGQHIKQLSRPASASIKIAPEPTPDSK
                     VRMVVITGPPEAQFKAQGRIYGLKEENFPGPKEEVKLETHIRVPASAGRVIGKGGK
                     TVNELQNLTAAAVVVPRDQTPDENDQVIVKIIGHFYASQMQRKIRDILAQVKQHQK
                     GQSNQAQARRK"

ORIGIN
        Query Match      65.1%;  Score 1112.4;  DB 10;  Length 2105;
        Best Local Similarity   86.2%;  Pred. No. 3.2e-269;
        Matches 1285;  Conservative   0;  Mismatches 178;  Indels 27;  Gaps  4;
```

```
JOURNAL   Mol. Cell. Biol. 17 (4), 2158-2165 (1997)
MEDLINE   97220007
PUBMED    9121465
REFERENCE 2 (bases 1 to 2021)
AUTHORS   Oleynikov,Y.S., Ross,A.F. and Singer,R.H.
TITLE     Direct Submission
JOURNAL   Submitted (24-SEP-1997) Cell Biology, Albert Einstein College of
          Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA
FEATURES            Location/Qualifiers
     source         1..2021
                    /organism="Gallus gallus"
                    /mol_type="mRNA"
                    /db_xref="taxon:9031"
                    /cell_type="fibroblast"
                    /dev_stage="embryo"
     gene           1..2021
                    /gene="ZBP1"
     CDS            129..1859
                    /gene="ZBP1"
                    /function="binds a zipcode element in beta-actin mRNA
                    3'UTR and is involved in its localization"
                    /note="beta-actin mRNA zipcode-binding protein"
                    /codon_start=1
                    /product="zipcode-binding protein"
                    /protein_id="AAB82295.1"
                    /db_xref="GI:2570921"
                    /translation="MNKLYIGNLNESVTPADLEKVPNDHKISFSGQFLVKSGYAFVDC
                    PDEQWAMKAIETFSGVEHLGKQLIEHSVPKKQRSRKIQIRNIPPQLRWEVLDGLLA
                    QYGTVENCEQVNTDSETAVVNVTYTNREQTRQAIMKLNGHQLENHVLKVSYIPDEQSV
                    QGPENGRRGFGARGAPRQGSPVTAGAPVKQQPVDIPLRLLVPTQYVGAIIGKEGATI
                    RNITKQTQSKIDVHRKENAGAEKAISIHSTPEGSAACKMLEIMQKEAKDTKTADE
                    VPLKILAHNFVGRLIGKEGRNLKKVEQDTETKITISSLQDLTLYNPERTITVKGSIE
                    NCCKAEQEIMKVRRAYENDVAAMSLQSHLIPGLNLAAVGLFPASSNAVPPPSVSG
                    AAPYSSFMPPEQETVHVFIPAQAVGAIIGKKGQHIKQLSRFASASIKIAPETPDSKV
                    RMVITGPPEAQFKAQGRIYGKLKEENFFGPKEEVKLETHIRVPASAAGRVIGKGGKT
                    VNELQNLTAAEVVVPRDQTPDENEQVIVKLIIGHFYASQMAQRKIRDILAQVKQQHQKG
                    QSGQIQARRK"
     misc_feature   138..155
                    /gene="ZBP1"
                    /note="encodes RNP2 RNA binding domain"
     misc_feature   236..260
                    /gene="ZBP1"
                    /note="encodes RNP1 RNA binding domain"
     misc_feature   354..374
                    /gene="ZBP1"
                    /note="encodes nuclear localization signal; NLS"
     misc_feature   621..632
                    /gene="ZBP1"
                    /note="encodes nuclear localization signal; NLS"
     misc_feature   747..773
                    /gene="ZBP1"
                    /note="encodes KH-domain #1"
     misc_feature   990..1016
                    /gene="ZBP1"
                    /note="encodes KH-domain #2"
     misc_feature   1062..1088
                    /gene="ZBP1"
                    /note="encodes REV-like nuclear export signal; NES"
     misc_feature   1374..1400
                    /gene="ZBP1"
                    /note="encodes KH-domain #3"
     misc_feature   1620..1646
                    /gene="ZBP1"
                    /note="encodes KH-domain #4"
ORIGIN

Query Match            53.6%;  Score 914.8;  DB 5;  Length 2021;
Best Local Similarity  81.3%;  Pred. No. 2.1e-219;
Matches 1089;  Conservative 0;  Mismatches 242;  Indels 9;  Gaps 2;

Qy      30 CCCGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCTGAAGGTCT 89
           |||||||||||||||||||||||||||  ||||| | |||||||||||| |||||||||
Db     523 CCAGGCAAGCCATCATGAAGCTGAATGAAGTTAAACGGGCACCAGTTGGAGAACCACGTGCTGTCAAAGTCT 582
```

```
Qy     690 AGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCTGGCCTGAACCTGGCTG 743
           ||||||| |||||||||||||||      ||||||||  |||| ||||| |||||||||
Db    1277 AGAATGACGTGGCTGCCATGAGCTTGCAGTCTCACCTCCACCTCATCCTGGTCTCAACCTGGCTG 1336

Qy     744 CTGTGTAGGTCTTTCCCAGCTTCATCCAGCGAGTCCGCCGCCTCCAGCAGCGTTACTG 803
           |||||| | ||| ||||||| || || || | || ||||| |||| ||||| ||||||
Db    1337 CTGTGTGGTCTCTTCCCAGTTCTTCTAGTGCCGGTCCGCCTCCTCCCCAGCAGGGTCACTG 1396

Qy     804 GGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTA 863
           ||||||||||||||| |||||| ||||||||||||||||||||||||||| ||||||| |
Db    1397 GGGCTGCTCCCTATGGCTCCTTCATGCAGGCTCCCGAGCAGGAGATGGTACAGGTGTTCA 1456

Qy     864 TCCCGCCCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGGCAGCACATCAAACAGCTCT 923
           |||| ||||||| ||||||||||||||||||||||||||  |||||||||||||||||||
Db    1457 TCCCCGCCCAGGCTGTGGGCGCCATCATCGGCAAGAAGGGCCAGCACATCAAACAGCTCT 1516

Qy     924 CCCGGTTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTC 983
           ||||||| ||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    1517 CCCGGTTCGCCAGCGCCTCCATCAAGATTGCTCCACCCGAAACACCTGACTCCAAAGTTC 1576

Qy     984 GTATGGTTATCATCACCTGACCCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATG 1043
           | ||||||  ||||||| |||||  |||||| || | ||||||||||||||||||||||||
Db    1577 GAATGGTTGTCATCACTGACCCCCAGAGGCTCAGTTCAAGGCTCAGGGAAGAATCTATG 1636

Qy    1044 GCAAACTCAAGGAGGAGAACTTCTTTGGTCCAAGAGGAGGAAGTGAAGCTGGAGACCCACA 1103
           |||||||  ||| |||||| ||||||||||||||||||||||||| ||||||||||||||||
Db    1637 GCAAACTAAAGAGGAGAATTTCTTTGGTCCAAGAGGAGGAAGTAAAGCTGGAGACCCACA 1696

Qy    1104 TACGTGTGCCAGCATCAGCAGCTGGCCGGTCATTGGCAGTGGCAAAGTGGAAAAACGGTGAACG 1163
           | ||||| ||||| |||||||||||||| |||||||||  ||||||||||  |||||||||| |
Db    1697 TCCGGGTCCCAGTTCAGCAGCTGGCCGGGTCATTGGCAAAGTGGCAAAACGGTGAATG 1756

Qy    1164 AGTTGCAGAATTTGACGGCAGCTGAGGTGTAGTGACCAGAGACCCCTGATGAGA 1223
           |||||||||| ||||| |||| |||||||  |  || || |||| ||| ||||||
Db    1757 AGTTGCAGAACTTGACCCCAGCCGAGGTGGTCGTGCCGAGACCAGGAGACCAGACGAGA 1816

Qy    1224 ACGACCAGGTCATCGTGAAAATCATCGACACATTCTATGCCAGTCAGATGGCTCAACGGA 1283
           |||||||||||||  |  || ||||||| |||||||||||| || |||||||| ||| |||
Db    1817 ACGACCAGGTCATTGTTAAGATCATCGGACATTCTATGCCCAGATGGCCCAGCGGA 1876

Qy    1284 AGATCCGAGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACCAGG 1343
           |||||||||||||||||||| || |||||||||| |||| |||||||||||||| |||||||
Db    1877 AGATCCGAGACATCCTGGCTCAAGTTAAGCAGCAACAGCACCAGAAGGGACAGAGCAACCAGG 1936

Qy    1344 CCCAGGCACGGAGGAAGTGA-CCAGCCCCTCCCTGTCCCTTNGAGTCCAGGACAACACG 1402
           |||||||||||| |||||||| ||||||||  |   || |||  | || |||||||||| |
Db    1937 CCCAGGCACGGAAGGAAGTGACCAGCCCCCTTTCCTGCTCCATTGGCTCCCAGATCAGCAGGGC 1996

Qy    1403 GGCAGA------------AATCGAGAGTGTGCTCTCCCCGGCAGGCCTGAGAA 1443
           ||||||            ||| || ||| ||| | ||  || ||||||||||||
Db    1997 CGAACACAGGACTGGAGGACGCGGTGAGGGCCCGGCCGTCTCCCTCCCTAGCAGGCCTGAGAA 2056

Qy    1444 TGAGTGGAATCCGGGACACNTGGCCCGGCTGTGAGATCAGGTTTGCCCA 1493
           ||||||||||| ||| | || || || || |||| |||||||||||| |||
Db    2057 TGAGTGGAATCAGGG-CGTTTGGGTCTGGCTGGGATCAGGTTTGCACA 2105
```

```
RESULT 15
AF026527
LOCUS       AF026527         2021 bp    mRNA    linear   VRT 30-OCT-1997
DEFINITION  Gallus gallus zipcode-binding protein (ZBP1) mRNA, complete cds.
ACCESSION   AF026527
VERSION     AF026527.1  GI:2570920
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 2021)
  AUTHORS   Ross,A.F., Oleynikov,Y., Kislauskis,E.H., Taneja,K.L. and
            Singer,R.H.
  TITLE     Characterization of a beta-actin mRNA zipcode-binding protein
```

```
Qy    90  CCTACATCCCCATGAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGCTTTG   149
          ||||||  ||| | |||||||   |||||| ||| || |||||||||||||| | ||||
Db   583  CCTACATCCCTGATGAGCAGTCCGTGCAGGGCCCAGAGAATGGGCGCGGGGTGGCTTTG   642

Qy   150  GCTCTCGGGGTCAGCCCCCGCCAGGGCTCACCTGTGGCAGCGGCCCAGCCAAGCAGC   209
          | | | |  | | || ||| ||||||| ||||| || |  ||  | | || ||||||
Db   643  GGGCCCGGAGGTGCCCCCCCGGCAGGGGCTCCCCTGTCACTGCGCAGGGGTCCAGTCAAGCAGC   702

Qy   210  AGCAAGTGACATCCCCTTCGGCTCCTCGGTGCCCACCAGTATGTGGGTGCCATTATTG   269
          |||  ||||| |||  | ||| || | || |||| | |||  |||| |||| || |||
Db   703  AGCCGTGTGACATCCCGCGCTCCGTCGGTGCCCAGCCAGTATGTGGGTGCCATCATTG   762

Qy   270  GCAAGGAGGGGCCACCATCCCGCCACCAGTCCAAGATAGACACCAGTCCAAGATAGACGTGC   329
          |||  |||||||||||| | |||||||||  | || || ||||  ||  |||||||  |||||
Db   763  GCAAGGAGGAGGGGCCACCATCACGCCACCAGTCCAAGCAGACAGTCCAAGATTGATGTGC   822

Qy   330  ATAGGAAGGACAAACGCAGGTGCAGCTTGTAGATGATCTTGGACATTATGCCATAAAGAGGCTAAGGACACCA   389
          || |||| |  |  | |||||| |||| ||   | ||| ||  |||| ||||| | | |||| || ||||||
Db   823  ACCGTAAGGAGGAGAATGCAGGAGCTGACGGTCTATCAGCATTCTTGGATCATGCCAGAAGAGGCAAAGGACACGA   882

Qy   390  GCTGCTCCTCCGCTTGTGAAGATCTTGGCAGATCCTGCCCCATAATAACTTTGTAGGGCGTTCTCA   449
          |||||||  |  |||| ||   | ||||| |||| | ||||| || || ||| | ||  || |||
Db   883  GCTGCTCTGCTGCTTGGCAAAATGATCTTTGAGGATCATGCCCCATAATCACTTTTGTGGGGCGCCTGA   942

Qy   450  AAACGGCTGACGAGGTTCCCCTGAAGATCCTGCCCCATAATAACTTTGTAGGGCGTTCTCA   509
          | ||| |||| || ||| ||  |||| | | ||| ||| || ||||   ||| ||||| |
Db   943  AGACAGCTGATGAAGTGCCTCTGAAAATCTTGCCCCATAATCAACACTTTGTGTGGGGCGCCTGA   1002

Qy   510  TTGGCAAGGAAGGACGGAACTTGAAGAAGGTAGAGACAAGGAGACAATACCACCA   569
          |||||||||| | | |||||  | |||||||| || | ||  || |||| |||||
Db  1003  TTGGCAAAGAAGGGCGGAACTTGAAGAAAGTGGAGCAGGATACTGAGACAAAAATCACCA   1062

Qy   570  TCTCCCTCGTTGCAAGAACCTTACCCTTTACAACCCTTTGAGAGGGGG   629
          ||||| | ||| || || ||| |||| | | ||  |||||  ||||||
Db  1063  TCTCATCCTTGGAGGACCTGACCCTGCTACACCCCGAAAAGAACGATCACAGTGCAGTGAAGGGCT   1122

Qy   630  CCATCGAGAATTGTTGCAGGGCCGAGCAGAGGGCAGAAATAATGAAGAAAGTTCGGGAGGCCTATG   689
          || |||||| ||||| ||   ||| ||| || |||| |   ||||||||  ||||  | ||||  |
Db  1123  CTATCGAGAACTGCTGCAAAGACAGGCAGCAGCAGGATCATGAGAAGTGAAGAAAGTGAGAGGCCTACG   1182

Qy   690  AGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGCTG   743
          || | || ||||| || ||||    ----||  |||||  | ||||||
Db  1183  AGAACGATGTTGCAGCCATGAGCCTGGTCGGTGCCATCATCTCATCCTGGCCTCAACCTGCTG   1242

Qy   744  CTGTAGGTCTTTTCCGAGCCTCATCCCGAGCTCATCCAGCGCGTTACTG   803
          ||| |||| | | |||| || ||||| | ||||| ||| |||| |  |
Db  1243  CAGTTGGCCTCTTCCCCTGCCTCCTCCCAGCAGTAACCTCCTCCCCAGCGACGTCTCCCG   1302

Qy   804  GGGCTGCTCCCATAGCTCCTTTATGCGAGCTCCCAGCGAGATGGTGAGGTGTTTA   863
          ||||||||||  ||| || |  ||||| || ||||||| ||| ||  |  ||||
Db  1303  GGGCTGCTCCAAGCTCCCATGAGC---CTCCGAGCGAGGCAGGACCGTACACGTCTTCA   1359

Qy   864  TCCCCGCCAGCAGTCGGCGCCCATCATCGGCGCAGAATTCAAGGCTCAGGAGAATCTATG   923
          ||||| ||||||| |  | || ||| || |||| || |||| |||| || ||||| | |
Db  1360  TCCCTGCCCAAGCGGTTGGTGCCATCATTGGCAAGAACGGGCCAGCATCAAGCAGCTCT   1419

Qy   924  CCCGGTTTGCCAGCGCCTTCCATCAAGATTGCACCACCGAAAACACCTGACTTCAAAGTTC   983
          ||||||| ||||| || || ||| ||| |  ||||||| |||| || || || ||||||
Db  1420  CCCGGTTCGCCAGCGCCCTTCTATTAAGATTGCCACCCCCGAAAGGGACTCCGAGCGTCCCG   1479

Qy   984  GTATGGTTATCATCACTGACACCGCCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATG   1043
          || ||||||||| || |||||||   |||| || | ||||  |  |||  || || | |
Db  1480  GCATGTGGTCATCACGCGCTCACCGGCCCCTCCAGAGCTCAGTTCAAGGCACAGGACGATTTATG   1539

Qy  1044  GCAAACTCAAGGAGGAGACATTCTTTGGTCCCAAGGAGGAGGTGAAGGCTGAGACCCCACA   1103
          ||||  || ||||||||| |  | || ||||||| |||||| |||||||| || |||||
Db  1540  GGAAGCTGAAGGAGGAGACTTCTTTGGGCCCAAGGAGGAGGTGAAGGTGGAGACGGTGAACG   1599

Qy  1104  TACGTGTGCCAGCATCAGCAGCTGGCGGTGCCCGGGTCATTGCCAAGGTGAAAAACGGTGAACG   1163
          | || ||||| || |||||||| || ||  || | || | ||||||| | ||||||| || |
Db  1600  TCCGTGTCCCCGCCTCGCTGCAGGGAGGTGATCGCAGCGGGTGATCCGCAAAAGGGGCAAAACCGTCAATG   1659

Qy  1164  AGTTGCAGAATTTGACGGCAGCTGAGGTGTGTAGTACCAAGAGACACCCCTGATGAGA   1223
          || |||||| | ||| || |||||||||| |||| |||| ||| |||||||||||||
Db  1660  AGCTGCAGAGACCTGCGCGGCTGCAGAGGTGCGTGGTTCCACGGGATCAGACCCCTGATGAGA   1719

Qy  1224  ACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGA   1283
          | || ||||||||||||||| ||||||| |||| ||||||||| |||||||| ||| |
Db  1720  ATGAGCAGGTCATTGTGAAGATCATCGGGCACTTCTATGCCAGCCAGATGGCACAGCGCA   1779

Qy  1284  AGATCCGAGACATCCTGGCCCAGGTTAAGCAGCATCAGAGGGACAGAGTAACCAGG   1343
          | ||||| ||||||||||||||||| |||||||| | || ||||| |  | ||||
Db  1780  AAATCCGGGACATCCTGGCCCAGGTGAAGCAGCAGCCCAGGGACAGGACAGAGGCGCCAGC   1839

Qy  1344  CCCAGGCACGGAGGAAGTGA   1363
          |||| ||||||||||| |||
Db  1840  TGCAAGCGGAGGAGGAAATGA   1859


Search completed: July 13, 2004, 20:23:55
Job time : 6733 secs
```

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      July 13, 2004, 20:24:00 ; Search time 792 Seconds
                          (without alignments)
                          10510.849 Million cell updates/sec

Title:       US-09-270-437D-5
Perfect score: 1708
Sequence:    1 agggacgcgtgcgcaccgcc........atttccctcaggttttaaaa 1708

Scoring table: IDENTITY_NUC
             Gapop 10.0 , Gapext 1.0

Searched:    318909 seqs, 2436941669 residues

Total number of hits satisfying chosen parameters:    6367818

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications NA:*
    1 : /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
    2 : /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
    3 : /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
    4 : /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
    5 : /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
    6 : /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
    7 : /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
    8 : /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
    9 : /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
   10 : /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
   11 : /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
   12 : /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
   13 : /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
   14 : /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
   15 : /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
   16 : /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
   17 : /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
   18 : /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
   19 : /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1703	99.7	1708	9	US-09-899-651-5	Sequence 5, Appli
2	1669.0	97.7	1946	9	US-09-899-651-7	Sequence 7, Appli
3	1283.6	75.2	1738	16	US-10-062-674-2082	Sequence 2082, Ap
4	1165	68.2	2224	9	US-09-873-637-1	Sequence 1, Appli
5	1085.4	63.5	1182	9	US-09-764-864-327	Sequence 327, App
6	1069.8	62.6	1129	9	US-09-764-864-745	Sequence 745, App
7	582.2	34.1	1740	9	US-09-735-705-347	Sequence 347, App
8	582.2	34.1	1740	9	US-09-850-716A-347	Sequence 347, App
9	582.2	34.1	1740	9	US-09-897-778-347	Sequence 347, App
10	582.2	34.1	1740	13	US-10-007-700-347	Sequence 347, App
11	582.2	34.1	1740	15	US-10-117-982-347	Sequence 347, App
12	582.2	34.1	1740	15	US-10-117-982-478	Sequence 478, App
13	582.2	34.1	1740	16	US-10-313-986-347	Sequence 347, App
14	582.2	34.1	1740	16	US-10-313-986-478	Sequence 478, App
15	582.2	34.1	1743	9	US-09-897-778-447	Sequence 447, App
16	582.2	34.1	1743	9	US-09-897-778-450	Sequence 450, App
17	582.2	34.1	1743	13	US-10-007-700-447	Sequence 447, App
18	582.2	34.1	1743	13	US-10-007-700-450	Sequence 450, App
19	582.2	34.1	1743	15	US-10-117-982-447	Sequence 447, App
20	582.2	34.1	1743	15	US-10-117-982-450	Sequence 450, App
21	582.2	34.1	1743	16	US-10-313-986-447	Sequence 447, App
22	582.2	34.1	1743	16	US-10-313-986-450	Sequence 450, App
23	582.2	34.1	1764	9	US-09-850-716A-428	Sequence 428, App
24	582.2	34.1	1764	9	US-09-897-778-428	Sequence 428, App
25	582.2	34.1	1764	13	US-10-007-700-428	Sequence 428, App
26	582.2	34.1	1764	15	US-10-117-982-428	Sequence 428, App
27	582.2	34.1	1764	16	US-10-313-986-428	Sequence 428, App
28	582.2	34.1	1799	16	US-10-313-986-485	Sequence 485, App
29	582.2	34.1	34555	15	US-10-117-982-479	Sequence 479, App
30	582.2	34.1	34555	16	US-10-313-986-479	Sequence 479, App
31	580.6	34.0	4159	9	US-09-899-651-4	Sequence 4, Appli
32	580.6	34.0	4181	9	US-09-735-705-175	Sequence 175, App
33	580.6	34.0	4181	9	US-09-954-456-715	Sequence 715, App
34	580.6	34.0	4181	9	US-09-850-716A-175	Sequence 175, App
35	580.6	34.0	4181	9	US-09-897-778-175	Sequence 175, App
36	580.6	34.0	4181	10	US-09-466-396A-175	Sequence 175, App
37	580.6	34.0	4181	13	US-10-007-700-175	Sequence 175, App
38	580.6	34.0	4181	15	US-10-117-982-175	Sequence 175, App
39	580.6	34.0	4181	16	US-10-313-986-175	Sequence 175, App
40	580.6	34.0	4434	13	US-10-116-802-145	Sequence 145, App
41	572.6	33.5	1740	15	US-10-117-982-483	Sequence 483, App
42	572.6	33.5	1740	16	US-10-313-986-483	Sequence 483, App
43	504	29.5	3283	9	US-09-899-651-8	Sequence 8, Appli
44	504	29.5	3667	15	US-10-097-940-146	Sequence 146, App
45	504	29.5	3667	17	US-10-648-593-48	Sequence 48, Appl

ALIGNMENTS

```
RESULT 1
US-09-899-651-5
; Sequence 5, Application US/09899651
; Patent No. US2002011470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-5

Query Match           99.7%;  Score 1703; DB 9; Length 1708;
Best Local Similarity 100.0%;  Pred. No. 0;
Matches 1708;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     1 AGGGACGCGCTGCCGCACCGCCCCCCAGTTTACCCCGGGGAGCCATCATGAAGCTGAATGGCCA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 AGGGACGCGCTGCCGCACCGCCCCCCAGTTTACCCCGGGGAGCCATCATGAAGCTGAATGGCCA 60

Qy    61 CCAGTTGGAGAACCATGCCCTGAAGGTCTCTACATCCCGATGAGCAGATAGCACAGGG 120
```

Db 1141 CAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTACC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   61 CCAGTTGGAGAACCATGGCCTCCTACATCCCCTGAAGGTCTCCTACATCCCCTGAAGGGG 120

Qy 1201 AAGAGACCAGACCCCTGATGAGACGACCAGGTCATCGTGAAAATCATCGGACATTTCTA 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 AAGAGACCAGACCAGACCCCTGATGAGACGACCAGGTCATCGTGAAAATCATCGGACATTTCTA 1260

Qy 1261 TGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAGCAGCA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 TGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAGCAGCA 1320

Qy 1321 TCAGAAGGACAGAGTAACCAGGCCCAGGACGCACGGAAGTGACCAGCCCCTCCTGTCC 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 TCAGAAGGACAGAGTAACCAGGCCCAGGACGCACGGAAGTGACCAGCCCCTCCTGTCC 1380

Qy 1381 CTTNGAGTCCAGACAACAACGGCAGAAATGCAGAGTGTGCTCTCCCCGGAGGCCTGA 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 CTTNGAGTCCAGACAACAACGGCAGAAATGCAGAGTGTGCTCTCCCCGGCAGGCCTGA 1440

Qy 1441 GAATGAGTGGGAATCCGGACACNTGGGCCGGGCTGTAGATCAGGTTGCCCACTTGATT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GAATGAGTGGGAATCCGGACACNTGGGCCGGGCTGTAGATCAGGTTTGCCCACTTGATT 1500

Qy 1501 GAGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCAAACACCCAATTGGC 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 GAGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCAAACACCCAATTGGC 1560

Qy 1561 CCAACACTGTNTGCCCCTCGGGGTGTCAGAAATTNTAGCGCAAGGCACTTTTAAACGTGG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 CCAACACTGTNTGCCCCTCGGGGTGTCAGAAATTNTAGCGCAAGGCACTTTTAAACGTGG 1620

Qy 1621 ATTGTTTAAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTGGGAGA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 ATTGTTTAAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTGGGAGA 1680

Qy 1681 AAAATAAAATTTCCTTCAGGTTTTAAAA 1708
          ||||||||||||||||||||
Db 1681 AAAATAAAATTTCCTTCAGGTTTTAAAA 1708

```
RESULT 2
US-09-899-651-7
; Sequence 7, Application US/09899651
; Patent No. US2002011470A1
; GENERAL INFORMATION:
;   APPLICANT: Chen, Yao-Tseng
;   APPLICANT: Gure, Ali
;   APPLICANT: Tsang, Solam
;   APPLICANT: Stockert, Elisabeth
;   APPLICANT: Jager, Elke
;   APPLICANT: Knuth, Alexander
;   APPLICANT: Old, Lloyd J.
;   TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
;   TITLE OF INVENTION: Associated
;   TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
;   FILE REFERENCE: LUD 5538
;   CURRENT APPLICATION NUMBER: US/09/899,651
;   CURRENT FILING DATE: 2001-07-06
;   PRIOR APPLICATION NUMBER: US/09/061,709
;   PRIOR FILING DATE: 1998-04-17
;   NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
;   LENGTH: 1946
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-899-651-7

Query Match       97.7%;     Score 1669.2;    DB 9;   Length 1946;
Best Local Similarity 99.8%;  Pred. No. 0;
Matches 1676;  Conservative 0;  Mismatches 3;  Indels 0;  Gaps 0;
```

Qy   30 CCCGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT 89
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   61 CCAGTTGGAGAACCATGGCCTCCTACATCCCCTGAAGGTCTCCTACATCCCCTGAAGGGG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   121 ACCTGAGAATGGGCGCCGAGGGGGCTTTGGCTCTCGGGGTCTCGGGGTCTCGGGGCTCACC 180

Db   121 ACCTGAGAATGGGCGCCGAGGGGGCTTTGGCTCTCGGGGTCTCGGGGTCTCGGGGCTCACC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181 TGTGGCAGCGCGGGGGCCCCAGCCAGCAGCAAGCAGCAGCAAGCAGCACATCCCCTCCTGGT 240

Db   181 TGTGGCAGCGCGGGGGCCCCAGCCAGCAGCAAGCAGCAGCAAGCAGCACATCCCCTCCTGGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   241 GCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGGCCACATCCCAACATCAC 300

Db   241 GCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGGCCACATCCCAACATCAC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   301 AAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAAACGCAGGTGCAGCTGAAAA 360

Db   301 AAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAAACGCAGGTGCAGCTGAAAA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   361 AGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTTGTAAGATGATCTTGGA 420

Db   361 AGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTTGTAAGATGATCTTGGA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   421 GATTATGCATAAAGAGAGCTAAGGACACCAAAAAACGGCTAAGGACGTTCCCTGAGATCT 480

Db   421 GATTATGCATAAAGAGAGCTAAGGACACCAAAAAACGGCTAAGGACGTTCCCTGAGATCT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   481 GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAAGGACGGAAGGT 540

Db   481 GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAAGGACGGAAGGT 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   541 AGAGCAAGATACCGAGACAAAAATCACCATCTCTGTTGCAAGACCTTACCCTTTACAA 600

Db   541 AGAGCAAGATACCGAGACAAAAATCACCATCTCTGTTGCAAGACCTTACCCTTTACAA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   601 CCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGA 660

Db   601 CCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGA 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   661 AATAATGAAGAAAGTTCGGGAGGCCATGAGAATGATGTGGCTGCCATGAGCTCTCACCT 720

Db   661 AATAATGAAGAAAGTTCGGGAGGCCATGAGAATGATGTGGCTGCCATGAGCTCTCACCT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   721 GATCCCTGGCCTGGCTCGTGTGCTGTGCTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCC 780

Db   721 GATCCCTGGCCTGGCTCGTGTGCTGTGCTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 GCCGCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGA 840

Db   781 GCCGCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   841 GCAGGAGATGGTGCAGGTGTTTATCCCCGGCCAGCCAGTCATCGGCAAGAA 900

Db   841 GCAGGAGATGGTGCAGGTGTTTATCCCCGGCCAGCCAGTCATCGGCAAGAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   901 GGGGCAGCAGCACATCAAACAGCTCTCCCCGGTTTGCCAGCGCCCTCCATCAAGATTGCACCACC 960

Db   901 GGGGCAGCAGCACATCAAACAGCTCTCCCCGGTTTGCCAGCGCCCTCCATCAAGATTGCACCACC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   961 CGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCACGCCAGAGGCCCAATT 1020

Db   961 CGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCACGCCAGAGGCCCAATT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1021 CAAGGCTCAGGGAAGAATCTATGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCCAAGGA 1080

Db 1021 CAAGGCTCAGGGAAGAATCTATGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCCAAGGA 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1081 GGAAGTGAAGCTGGAGACCCAGAATCAGCGTGTGCCAGCATCAGCGTGCCAGCTGGCCGGGTCATTGG 1140

Db 1081 GGAAGTGAAGCTGGAGACCCAGAATCAGCGTGTGCCAGCATCAGCGTGCCAGCTGGCCGGGTCATTGG 1140

```
Db   268  CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT  327

Qy    90  CCTACATCCCGATGAGCAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGCTTTG  149
          ||||||||||||||||| ||||||||||||||||| ||||||||||||||||||||||||
Db   328  CCTACATCCCGATGAGCAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGCTTTG  387

Qy   150  GCTCTCGGGGTCAGCCCCGCCGGGCTCACCTGTGGCGCCCCCAGCCAGCAGC          209
          |||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   388  GCTCTCGGGGTCAGCCCCGCCCGGGCTCACCTGTGGCAGCCCCAGCCAGCAGC          447

Qy   210  AGCAAGTGACATCCTCGGTTCTTCGGCTCCTGGTGCTGCCACCCAGTATGTGGGTGCCATTATTG  269
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   448  AGCAAGTGACATCCTCGGTTCTTCGGCTCCTGGTGCTGCCACCCAGTATGTGGGTGCCATTATTG  507

Qy   270  GCAAGGAGGGGGCCACCATCCGCAACATCACAAAAACAGACCCAGTCCAAGATAGACGTGC  329
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508  GCAAGGAGGGGGCCACCATCCGCAACATCACAAAAACAGACCCAGTCCAAGATAGACGTGC  567

Qy   330  ATAGGAAGGAGAACGCAGTGCAGCTGAAGATGTTGGGATTATGCATAAAGAGGCTAGGACACCA  389
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568  ATAGGAAGGAGAACGCAGTGCAGCTGAAGATGTTGGGATTATGCATAAAGAGGCTAGGACACCA  627

Qy   390  GCTGCTCCTCCGCTTGTGAAGATGATCTTGGAGATCTCTGAAGATCCTGCCCCATAATAACTTTGTAGGGCGTCTCA  449
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628  GCTGCTCCTCCGCTTGTGAAGATGATCTTGGAGATCTCTGAAGATCCTGCCCCATAATAACTTTGTAGGGCGTCTCA  687

Qy   450  AAACGGCTGACGAGGTTCCCCTGCCCATAGAGTACGGAACCTGAAGACCTTTAAACGTGGATTGTTTAAAGAAGCACTTTAAACGTGGATTGTTTAAAGAACACCA  509
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688  AAACGGCTGACGAGGTTCCCCTGCCCATAGAGTACGGAACCTGAAGACCTTTAAACGTGGATTGTTTAAAGAAGCACTTTAAACGTGGATTGTTTAAAGAACACCA  747

Qy   510  TTGGCAAGGAAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAAATCACCA  569
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748  TTGGCAAGGAAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAAATCACCA  807

Qy   570  TCTCCTCGTTGCAAGACCCTTACAACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGG  629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   808  TCTCCTCGTTGCAAGACCCTTACAACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGG  867

Qy   630  CCATCGCAGAATTGTTGCAGGGCCGAGCAGCAGGAATAATGAAGAAAGTTCGGGAGGCCTATG  689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   868  CCATCGAGAATTGTTGCAGGGCCGAGCAGCAGGAATAATGAAGAAAGTTCGGGAGGCCTATG  927

Qy   690  AGAATGATGTGGCGCCATGAGAGCTCTCACCTGAGATTCCCTGGCCTGAACCTGGCTGCTGTAG  749
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   928  AGAATGATGTGGCGCCATGAGAGCTCTCACCTGAGATTCCCTGGCCTGAACCTGGCTGCTGTAG  987

Qy   750  GTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCGCTTACTCGGGGCTG  809
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   988  GTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCGCTTACTGGGGGCTG  1047

Qy   810  CTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGTGTTTATCCCG  869
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1048  CTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGTGTTTATCCCG  1107

Qy   870  CCCAGGCAGTGGGCGCCATCATCGGCAGAAGGGGCAGCACATCAAACAGCTCTCCCGGT  929
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1108  CCCAGGCAGTGGGCGCCATCATCGGCAGAAGGGGCAGCACATCAAACAGCTCTCCCGGT  1167

Qy   930  TTGCCAGCGCCTCCATCAAGAGATTGCACCAGCACCACACCTGACTCGACTCGTATCGTATGG  989
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1168  TTGCCAGCGCCTCCATCAAGAGATTGCACCAGCACCACACCTGACTCGACTCGTATGG  1227

Qy   990  TTATCATCACTGACACCGGAGAACTTCTTTGGTTCCCAAGGAGGAAGATCCATGATGCAAAC  1049
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1228  TTATCATCACTGACACCGGAGAACTTCTTTGGTTCCCAAGGAGGAAGAATCCATGATGCAAAC  1287

Qy  1050  TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGAAGTGAAGCTGGAGACCCACATACGTG  1109
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1288  TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGAAGTGAAGCTGGAGACCCACATACGTG  1347

Qy  1110  TGGCCAGCATCAGCAGCTGCCGGGTCATTGGCAAAGGTGGAAAAACGTGACGAGTTGC  1169
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1348  TGGCCAGCATCAGCAGCTGCCGGGTCATTGGCAAAGGTGGAAAAACGTGACGAGTTGC  1407
```

```
Db  1170  AGAATTGACGGCAGCTGAGGTGTGAGTAGTACCAAGAGACCAAGAGACACCCTGATGAGAACGACC  1229

Qy  1408  AGAATTGACGGCAGCTGAGGTGTGAGTAGTACCAAGAGACCAAGAGACACCCTGATGAGAACGACC  1467

Qy  1230  AGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGTGGCTCAACGGAGATCC  1289
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1468  AGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGTGGCTCAACGGAGATCC  1527

Qy  1290  GAGACATCCTGGCCCAGGTTAAGCACGAGCATCAGAAGGACAGAGTAACCAGGCCCAGG  1349
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1528  GAGACATCCTGGCCCAGGTTAAGCACGAGCATCAGAAGGACAGAGTAACCAGGCCCAGG  1587

Qy  1350  CACGGAGGAAGTGACCAGCCCCTCCTGTCCCTTNGAGTCCAGGACCAACAACACGGGGCAGAA  1409
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1588  CACGGAGGAAGTGACCAGCCCCTCCTGTCCCTTNGAGTCCAGGACAACAACACGGGGCAGAA  1647

Qy  1410  ATCGAGAGTGTGTCTCCCGGCAGGCTGAGAATGAGTGGGAATCCGGGACACNTGGGC  1469
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1648  ATCGAGAGTGTGTCTCCCGGCAGGCTGAGAATGAGTGGGAATCCGGGACACNTGGGC  1707

Qy  1470  CGGGCTGTAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGA  1529
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1708  CGGGCTGTAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGA  1767

Qy  1530  TCTNTCAGCCCCAAACACCCACCCAATTGGCCCAATTGGCCCAAACACTGTNTGCCCCTCGGGGTGTCAG  1589
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1768  TCTNTCAGCCCCAAACACCCACCCAATTGGCCCAATTGGCCCAAACACTGTNTGCCCCTCGGGGTGTCAG  1827

Qy  1590  AAATTNTAGCGCAAGGCACTTTAAACGTGGATTGTTTAAAGAAGCTCTCCAGGCCCCAC  1649
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1828  AAATTNTAGCGCAAGGCACTTTAAACGTGGATTGTTTAAAGAAGCTCTCCAGGCCCCCAC  1887

Qy  1650  CAAGAGGGTGGATCACACCTCAGTGGGAAGAAAAAATAAAATTTCCTTCAGGTTTTAAAA  1708

Db  1888  CAAGAGGGTGGATCACACCTCAGTGGGAAGAAAAAATAAAATTTCCTTCAGGTTTTAAAA  1946
```

RESULT 3
US-10-062-674-2082/c
; Sequence 2082, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 2082
; LENGTH: 1738
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2004000559A1 457674.1
US-10-062-674-2082

```
Query Match    75.2%;   Score 1283.6;   DB 16;   Length 1738;
Best Local Similarity  93.6%;   Pred. No. 0;
Matches 1615;   Conservative 0;   Mismatches 64;   Indels 46;   Gaps 25;

Qy    30  CCCGGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT  89
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1735  CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT  1676

Qy    90  CCTACATCCCCGATGAGCAGCAGATAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGGCTTTG  149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1675  CCTACATCCCCGATGAGCAGCAGATAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGGCTTTG  1616
```

RESULT 4
US-09-873-637-1
; Sequence 1, Application US/09873637
; Patent No. US2002061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2224
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-873-637-1

Query Match   68.2%; Score 1165; DB 9; Length 2224;
Best Local Similarity   82.8%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 260; Indels 33; Gaps 6;

Db   705 AGCCAGTTGGACATCCCTCTCCGGCTCCTGGTGCCTACGCAGTGATGTGAGGCGCTATCATTG 764
Qy   270 GCAAGGAGGGGGCACCATCCGACAACTCCGAAAATCACAAAACAGACCCAGTCAAGATAGACGTGC 329
         |||||||||| |||||||||||| ||| ||||||||||||| |||||||||| |||||||||||||
Db   765 GCAAGGAGGGGTGCCACCATCCGAAAATCACAAATCACAAAACAGACGCAGTCAAAATAGACGTGC 824
Qy   330 ATAGGAAGGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGG 389
         ||||||||||||| || || || ||||||||||||||||||||||| |||||| |||| ||||
Db   825 ATAGGAAGGAGAATGCGGGCGTGCGGGAGAGGCCATCAGCGTGCATTCAACCCCTGAAG 884
Qy   390 GCTGCTCTCCGCCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 449
         |||||||||| |||||||||||| |||||||||||||| ||||||||||||||| |||||
Db   885 GCTGCTCCTCCCGCGTGCAAGATGATGATCTTTGGAGATTATGCACAAGGAGGCAAAGGACACCA 944
Qy   450 AAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGGGTCTCA 509
         ||||| ||||||||||||||||||||| |||||||||||||||||||||||||||| |||
Db   945 AAACGGCAGATGAAGTTCCCTGAAGATCCTGGCTCATCATAACAACTTCGTCGGGGACTCA 1004
Qy   510 TTGGCAAGGAGACGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAATCACCA 569
         ||||||||||||||||||| |||||||||||| |||| |||| |||||| |||||||
Db 1005 TTGGCAAGGAGGCCGAACCTGAAGAAGGTGAGCGAGCAGAGCGAAGATCACCA 1064
Qy   570 TCTCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGGAGGACCATCACTGTGAAGGGGG 629
         |||||||||||||||||| |||||| |||||| |||||||||| |||||||||||||||||
Db 1065 TCTCATCGTTGCAAGACCTCAGGCTCTATAACCCTGAGGAGGACCATCACTGTGAAGGGCG 1124
Qy   630 CCATCGAGAATTGTTGCAGGGGCCGAGCAGGAGAAATAATGAAGAAAGTTCGGAGGGCCTATG 689
         |||| ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db 1125 CCATTGAGAACTGTTGCAGGGGCCGAGCAGCAGGAGATCATGAAGAAAGTTCGAGAGGGCTTACG 1184
Qy   690 AGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCTGGCCTGAACCTGGCTG 743
         ||||||||||| ||||||||||| |||||||||| ||||| |||||||||||||
Db 1185 AGAACGACGTGGCCGTGGCCATGAGCTTGCAGTCCACCTCATCCCTGGGCTTAACCTGCTG 1244
Qy   744 CTGTAGGTCTTTTCCCAGCTTCATCCAGCGAGTCCCGCCTCCCAGCAGCGTTACTG 803
         |||||||||||| |||||||||||||| ||| |||||| ||||||||||||||||||
Db 1245 CTGTAGGTCTCTTCCCAGCTTCATCCAGCGAGCCGCTGCTCCCTCCCAGCAGTGTCACTG 1304
Qy   804 GGGCTGCTGTCCCTATAGCTCCTTTATGCAGGCTCCCCGAGCAGGAGATGGTGTCAGGTGTTA 863
         ||||||||| ||||||||||| |||||||| ||||| ||||||||||||||| ||| |||||||
Db 1305 GGGCTGCTGTCCCTATAGCTCCTTCCTGCAGCAGGCCGCTGGCAGGAGATGGTGTACAAGTGTTCA 1364
Qy   864 TCCCCGCCAGCAGTGGGCGCCATCATCTGGCGCAAGAGAGGGGCGCACATCAAACAGCTCT 923
         ||||||||||||||| ||| |||| ||| |||||||||| ||||||||| ||||||||||||
Db 1365 TCCCCGCCAGCGTGTGGGCGCCATCATTGGCAAGAGAGAGGGGCAGCACATCAAACAACTCT 1424
Qy   924 CCCGGTTTGCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTC 983
         |||| ||| |||| |||||||||| |||||||||||| ||||||||||||||||||||||
Db 1425 CCCGTTTCGCAGCGCCTCCATCAAGATTGCTCCACCAGAAACACCTGACTCCAAAGTTC 1484
Qy   984 GTATGGTTATCATCACTGGACCGGCCCAGAGGCCCCAGAGGCTCAGGGAAGAATCTATG 1043
         |||| |||||||||||| |||| |||||| ||| ||||| ||||||| |||||||||||||
Db 1485 GAATGGTCGTCATCACTGGACCCGCCCAGAGGCTCAGTTCGCAGGGAAGAATTTATG 1544
Qy 1044 GCAAACTCAAGGAGGAGACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCCACA 1103
         ||||| |||||||||||||| ||||| |||||||||||||||| ||||| |||||||||||
Db 1545 GCAAACTAAAGGAGGAGAATTCTTTGGTCCCAAGGAGGAAGTGAAAGCTAGAGACCCCACA 1604
Qy 1104 TACGTGTGCCAGCATCAGCTGCAGCTGCGCCCGGTCATTGCAAAGTGAAAAACCGGTGAACG 1163
         ||| || ||||| ||||||||| |||||| |||||||||| |||||||||| ||||||||||
Db 1605 TACGGGTTCGGCTTCAGCAGCCCGTCATCGGCAAGTGAAGCGGCAAAACGGTGAATG 1664
Qy 1164 AGTTCGAGAATTTGACGCGGCAGCTGAGTGGTAGTACCAAGACCAGCAGCCCTGATGAGA 1223
         |||||||||| ||| || |||||||| ||||| ||||||||||||||||| |||||||||||
Db 1665 AGTTCGCAGAATTTGACTGCAGCTGAGTGGTGTGCCAAGACCAGCAGCCCCGGATGAGA 1724
Qy 1224 ACGACCAGGTCATCGTGAAAATCATCGGACATTTCATGCCAGTTCAGATGGCTCAACGGA 1283
         ||||||| |||| ||| ||||||||| |||| ||||||| ||||||||||||||||||||
Db 1725 ACGACCAAGTCATTGTTAAGATCATCGGACAATTTCATGCCAGTTCAGATGGCTCAGCGGGA 1784
Qy 1284 AGATCCGAGACATCCTGCCCAGTTAAGCAGCAGAGGACCAGAGTAACCAGG 1343
         ||||||||||||||||| |||||||||||||| || |||||| ||||||||
Db 1785 AGATCCGAGACATCCTGCTCGGCTCAAGTTAAGCAACGACAGCACCAGGCAACGCAACCTGG 1844

Qy 1344 CCCAGGCACGGAGGGAGTGA-CCAGCCCTCCTCCCTGTCCCTTGAGTGAGTCCAGGACAACAACG 1402
         |||||||||||||||||||| ||||||||| |||||||||||| | ||||||||||||||||||
Db 1845 CCCAGGCACGGAGGGAGTGACCATTGGGCCTCCGCCCCTCCTGTCCCATTGGCTCTCCCGGCAGGCCTGAGG 1904
Qy 1403 GGCAGAA--------------------ATCGAGAGTGTGCTCTCCCCGGCAGGCCTGAG 1441
         ||||||| | |||||| || ||| |||| ||||||||||||
Db 1905 AGGAACACAGAACTGGAGGGGCGGGGTGGAGGGCCGGTGTGTGTTTTTCCCAGGCCTGAG 1964
Qy 1442 AATGAGTGGAATCCCGGACACNTGGCCGGCCTGTAGATCCAGGTTGCCCACTTGATTG 1501
         ||||||||| ||| | ||| ||||| ||| ||||||||||| |||||| ||| |||
Db 1965 AATGAGTGGGATCAGGG-CATTTGGGCGTCCTGCGTCGGAGATCAGGTTTGCCACACTGTATTG 2023
Qy 1502 AGAAAGAGTTCCAGTGAGGAACCCTGATCNTCAGCCCCAAAACACCCAATTGGCC 1561
         ||||| ||||||||| |||||| |||||| ||| ||||||||| ||||| |||
Db 2024 AGAACAATGTTCCAGTGAGGAATCCTGATCTTCTCGCCCCCAA--TTGAGCCAGCTGCCCA 2081
Qy 1562 CAACACTGTNTGCCCCTCGGGGTGTCAGAATTNAGCCGAAGGCACTTTAAACGTGGA 1621
         | ||||||| |||| |||||| | ||||||| ||| | |||||||||||| ||||
Db 2082 CAGCCCCACCCCTGAATATCACCATTGCAATCATCATGCAATAGCTTGGGTTGCTTTTAAACGTGGA 2141
Qy 1622 TTGTTAAAGAAGCTCTCCAGGACCACTGAAGGGGG 1681
         |||| |||||||||| |||||||||||||||||||
Db 2142 TTGTCT--TGAAGTTCTCCAGCCTCCATGGAAGGATGGGTCAGATCCCAGTGGGGAAGAG 2199
Qy 1682 AAATAAAATTTCCTTCAGGTTTTA 1705
         |||||||||||||||||||||||||
Db 2200 AAATAAAATTTCCTTCAGGTTTTA 2223

RESULT 5
US-09-764-864-327
; Sequence 327, Application US/09764864
; Patent No. US2002013275A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 327
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-327

Query Match   63.5%; Score 1085.4; DB 9; Length 1182;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1102; Conservative 6; Mismatches 0; Indels 6; Gaps 1;

Qy   601 CCCTGAGGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTGTTGCAGGGGCCGAGCAGGA 660
         |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   18 CGCTGAGGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTGTTGCAGGGGCCGAGCAGGA 77
Qy   661 AATAATGAAGAAAGTTCGGAGGGCCTATGAGAATGATGTGGCTGCCATGAGC-----TC 714
         |||||||||||||||||||||||| |||||||||| ||| |||||||||||| ||
Db   78 AATAATGAAGAAAGTTCGGAGGGCCTATGAGAATGAGATGTGGCTGCCATGAGCTGCAGTC 137
Qy   715 TCACCTGATCCTGGCCTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCAGGCGC 774
         ||||||||||||||||||||| ||||||||||||||| ||||||| |||||||||||||||
Db   138 TCACCTGATCCTGGCCTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGC 197
Qy   775 AGTCCCGCCTCCCAGCAGCGCTGCTGCTACTGGGCGTCCTGCTCCTTATGCAGGC 834
         ||||||||||||||||||||||| ||||||||||||| ||||||||||||||||||
Db   198 AGTCCCGCCTCCCAGCAGCAGCGCTGCTGCTACTGGGCGTCCTGCTCCTTATGCAGGC 257
Qy   835 TCCCGAGCAGGAGATGGTGCAGTGTTTATCCCCGCCAGCAGTCGGCCCATCATCGG 894
         ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   258 TCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGCCAGCAGTGGGCGCCATCATCGG 317

```
; ORGANISM: Homo sapiens
US-09-764-864-745

Query Match            62.6%;   Score 1069.8;   DB 9;   Length 1129;
Best Local Similarity  98.7%;   Pred. No. 0;
Matches 1084;   Conservative  3;   Mismatches  3;   Indels  5;   Gaps  1;

Qy  617 ACTGTGAAGGGGCCATCGAAGAATTGTTGCAGGGCCGAGCAGGAAATAATGAGAAAGT 676
        |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db    1 ACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGAAATAATGAGAAAGT 60

Qy  677 CGGGAGGGCCTATGAGAATGATGTGCTGCCATGAGC------TTCTCACCTGATCCTGGC 730
        |||||||||||||||||||||||||||||||||||||      |||||||||||||||||
Db   61 CGGGAGGGCCTATGAGAATGATGTGCTGCCATGAGCCTGCAGTTCTCACCTGATCCTGGC 120

Qy  731 CTGAACCTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCGCGGCCTCCC 790
        ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  121 CTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCGCGGCCTCCC 180

Qy  791 AGCAGCGTACTGGGCGTGCTCCCATAGCTCCTTATGCAGGCTCCCGAGCAGGAGATG 850
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AGCAGCGTACTGGGCGTGCTCCCATAGCTCCTTATGCAGGCTCCCGAGCAGGAGATG 240

Qy  851 GTGCAGGTGTTTATCCCCGGTTGCCAGCGCGCCTCATCAAGATTGACCACCGAAACACCT 910
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 GTGCAGGTGTTTATCCCCGGTTGCCAGCGCGCCTCATCAAGATTGACCACCGAAACACCT 300

Qy  911 ATCAAACAGCTCTCCCGGTTGTATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAG 970
        |||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db  301 ATCAAACAGCTCTCCCGGTTGCCAGCGCCTCATCAAGATTGACCGCCAGAGGCCCAATTCAAGGCTCAG 360

Qy  971 GACTCCAAAGTTCGTATGGTTATCATCACTGGACGAGGAGAACTTCTTTGGTCCCAAGGAGAAGTGAAG 1030
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GACTCCAAAGTTCGTATGGKTATCATYACTGGACGAGGAGAACTTCTTTGGTCCCAAGGAGAAGTGAAG 420

Qy 1031 GGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGAAGTGAAG 1090
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGAAGTGAAG 480

Qy 1091 CTGGAGACCCACATACGTGTGCCAGCATCAGCAGTCGGGGTCATTGGCAAAGGTGA 1150
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CTGGAGACCCACATACGTGTGCCAGCATCAGCAGTCGGGGTCATTGGCAAAGGTGA 540

Qy 1151 AAAACGGTGAACGAGTTGACAGAATTGACGGCAGTGAGTGGTAGTACCAAGAGACCAG 1210
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 AAAACGGTGAACGAGTTGACAGAATTTGACGGCAGTGAGTGGTAGTACCAAGAGACCAG 600

Qy 1211 ACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAG 1270
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAG 660

Qy 1271 ATGGCTCAACGGCGAAGATCCGAGACATCCGGACGTTAAGCAGCATCAGAAGGGA 1330
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ATGGCTCAACGGCGAAGATCCGAGACATCCGGACGTTAAGCAGCATCAGAAGGGA 720

Qy 1331 CAGAGTAACCAGGCCCAGGCCACGGAGGAAGTGACCAGCCCTCCCGTCCCTTNGAGTCC 1390
        ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db  721 CAGAGTAACCAGGCCCAGGCCACGGAGGAAGTGACCAGCCCTCCCGTCCCTTCGAGTCC 780

Qy 1391 AGGACAACAACGGACACTGAGAATCGAGAGTGCTCTCCCCGGCCACTTGAGAATGAGTGG 1450
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 AGGACAACAACGGACACTGAGAATCGAGAGTGCTCTCCCCGGCCACTTGAGAATGAGTGG 840

Qy 1451 GAATCCGGGACACNTGGGCCCGGGCGTGTAGATCAGGTTTGCCCACTTGAGAAGAATG 1510
        |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db  841 GAATCCGGGACACCTGGGCCCGGGCGTGTAGATCAGGTTTGCCCACTTGAGAAGAATG 900

Qy 1511 TTCCAGTGAGGAACCCTGATCNTCAGCCCAAACACCCAACTGCCCAACACTGT 1570
        ||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  901 TTCCAGTGAGGAACCCTGATCTTCAGCCCAAACACCCAACTGCCCAACACTGT 960

Qy 1571 NTGCCCCTCGGGGTGTCAGAAATTNTAGCGCAAGGCACTTTAAACGTCGGATTGTTTAAA 1630
        |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
```

```
Qy  895 CAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGTTTGCCAGGGCCTTCCATCAGAGATTGC 954
        ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  318 CAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGTTTGCCAGCGCCTTCCATCAGAGATTGC 377

Qy  955 ACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGC 1014
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  378 ACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGC 437

Qy 1015 CCAATTCAAGGCTCAGGAAGAGATCTATGCAACTCAAGGAGGAGAACTTCTTTTGGTCC 1074
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  438 CCAATTCAAGGCTCAGGAAGAGATCTATGCAACTCAAGGAGGAGAACTTCTTTTGGTCC 497

Qy 1075 CAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGTCGGCCGGGGT 1134
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  498 CAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGTCGGCCGGGGT 557

Qy 1135 CATTGGCAAAGTGGAAAAAACGGTGAACGAGTTGACGACGGCAGTGAGTGGT 1194
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  558 CATTGGCAAAGTGGAAAAAACGGTGAACGAGTTGACGACGGCAGTGAGTGGT 617

Qy 1195 AGTACCAAGAGACCAGACCCCTGATGAGAACGAGTCATCGTGAAAATCATCGGACA 1254
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  618 AGTACCAAGAGACCAGACCCCTGATGAGAACGAGTCATCGTGAAAATCATCGGACA 677

Qy 1255 TTTCTATGCCAGTCAGATGGCTCAACGGCGAAGATCCGAGACATCCTGGCCCAGGTTAAGCA 1314
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  678 TTTCTATGCCAGTCAGATGGCTCAACGGCGAAGATCCGAGACATCCTGGCCCAGGTTAAGCA 737

Qy 1315 GCAGCATCAGAGGAGGACAGAGTAACCAGGCCCAGGCCACGGAGGAAGTGACCAGCCCTCC 1374
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  738 GCAGCATCAGAGGAGGACAGAGTAACCAGGCCCAGGCCACGGAGGAAGTGACCAGCCCTCC 797

Qy 1375 CTGTCCCTTNGAGTCCAGGACAACAACGGGCAGAAATCGAGAGTGTGCTCTCCCCCGCAG 1434
        |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  798 CTGTCCCTTCGAGTCCAGGACAACAACGGGCAGAAATCGAGAGTGTGCTCTCCCCCGCAG 857

Qy 1435 GCCTGAAATGAGTGGGAATCCGGGACACNTGGGCCGGGTTGTAGATCAGGTTGCCCAC 1494
        |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  858 GCCTGAAATGAGTGGGAATCCGGGACACCTGGGCCGGGTTGTAGATCAGGTTGCCCAC 917

Qy 1495 TTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCACCCA 1554
        |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db  918 TTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGATCTCTCAGCCCCAAACACCCACCCA 977

Qy 1555 ATTGGCCCAACACTGTNTGCCCCTCGGGGTGTCAGAAAATTNTAGCGCAAGGCACTTTTAA 1614
        |||||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||||
Db  978 ATTGGCCCAACACTGTCTGCCCCTCGGGGTGTCAGAAAATTCTAGCGCAAGGCACTTTTAA 1037

Qy 1615 ACGTGGATTGTTTAAAGAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGATCACACCTCAGTG 1674
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1038 ACGTGGATTGTTTAAAGAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGATCACACCTCAGTG 1097

Qy 1675 GGAAGAAAATAAAATTTCCTTCAGGTTTTAAAA 1708
        |||||||||||||||||||||||||||||||||
Db 1098 GGAAGAAAATAAAATTTCCTTCAGGTTTTAAAA 1131
```

Db   961 CTGCCCCTCGGGTGTGTCAGAAATTCTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1631 GAAGCTCTTCCAGGCCCCACCCAGAGAGGGTGGATCACACCCTCAGTGGGAAGAAAAATAAAAT 1690
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 GAAGCTCTTCCAGGCCCCACCCAGAGAGGGTGGATCACACCCTCAGTGGGAAGAAAAATAAAAT 1080

Qy 1691 TTCCTTCAGGTTTTAAAA 1708
          ||||||||||||||||||
Db 1081 TTCCTTCAGGTTTTAAAA 1098

RESULT 7
US-09-735-705-347
; Sequence 347, Application US/09735705
; Patent No. US20020052329A1
GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-705-347

Query Match   34.1%;   Score 582.2;   DB 9;   Length 1740;
Best Local Similarity   66.2%;   Pred. No. 6.9e-171;
Matches 858;   Conservative   0;   Mismatches 433;   Indels   6;   Gaps   1;

Qy   37 AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT 96
         ||| || ||||||| || || |||||| || |||||
Db 402 AGCACTAGACAAACTGAATGGATTGAGATTTCAGTTAGAGAATTCACCTTGAAGTAGCCTATAT 461

Qy   97 CCCCGATAGCAGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGCTTTGGCTCTCG 156
         ||| ||| | | | |||||| ||||| ||| ||||||| || |||||| |
Db 462 CCCTGATGAGAAACCCTTGCAGCAGCCCGGGGGCCCCAGCAGCAGCCAAGT 521

Qy 157 GGGTCAGCCGCCCGCCGGGGCTCACCTGCTGGGGCGCAAAACCCTTGCAGCAGCCAGCAGT 216
         || || ||| | | ||| || | |
Db 522 TGGGCAGGAGGGCTCCTCAAGGCAGGGGGTCTCCAGGATCCGTATCCAAGCAGCAGAAACCATG 581

Qy 217 GGACATCCCCCTTCGGCTCCTGCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCCAAGGA 276
         |||| || || ||| | || ||||| || |||| ||| |||||| || |||
Db 642 AGGTGCCACCATTCCGAACATCACCAAACAGACCAGTCTAAAATGCGATGTCACCGTAA 701

Qy 277 GGGGGCCCACCATCCGCAACATCACCAAAACAGACACCAGTCCAAGTAGACGTGCATAGGAA 336

Qy 337 GGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTTGCACTCCACCCCTGAGGGCTGCTC 396
         || || || ||| || |||| | | || || | |
Db 702 AGAAAATGCGGGGGGGGCTGCTGAGAAGTCGATTACTATATCCCTCTCTACTCCTGAAGGCCACCTC 761

Qy 397 CTCCGGCTTGTAAGATGATCTTGGAGATTATGCCATAAAAGAGGGCTAAGGACACCCAAAAACGGC 456
         | |||| | | || | | ||| || |||| || | || |
Db 762 TGCGGCTTGTAAGTCTATTCTCGAGATTATGCATAAGGAGGCTCAAGAGATATAAAATTCAC 821

Db   457 TGACGAGGTTCCCCTGAAGATCCTGCGCCCATAATAACTTTGTAGGCGTCTCATTGGCAA 516
          ||||||||||||||||||||| |||||||||||||||||||| ||||| |||||||
Qy   822 AGAAGAGATCCCCTGAAGATTTTAGCTCATAATAACTTTGTTGACGTCTTATTGGTAA 881

Db   517 GGAAGGACGGAACCTGAAGAAGTAGAGCAAGTAGAGCAAGATACCGAGACAAAATCACCATCTCCTC 576
          ||||||| ||| || ||| ||||||||| ||| ||| ||||||| |||||||
Qy   882 AGAAGGAAGAAATCTTAAAAAAATTGAGCAAGTAGAAATTCAGACACTAAAATCACGATATCTCC 941

Db   577 GTTGCAAGACCTTACCCTTACAACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA 636
          ||||||| | |||| |||||||||| || || ||| ||| | ||
Qy   942 ATTGCAGGAATTGACGCTGTATAATCCAGAACGCACTATTACAGTTAAAGCAATGTTGA 1001

Db   637 GAATTGTTGCAGGGCCGAGCAGGAGAATAATGAAGAAGTTCGGAGGCCTATGAGAATGA 696
          || |||||| || || |||||| || ||||||||||| ||| ||||| |||||
Qy 1002 GACATGTGCCAAAGTGAGGAGGAGATCATGAAGAAAATCAGGGAGTCTTATGAAAATGA 1061

Db   697 TGTGGCTGCCATGA------GCTCTCACCTGATCCTGGCCTGAACCTGGCTGCTGTGTAGG 750
          | |||||| | || ||| || | || || | |||| |||| |||
Qy 1062 TATTGCTTCTATGAAATCTTCAAGCACATTTAATTCCTGGATTAAATCTGAACGCCCTTGGG 1121

Db   751 TCTTTTCCCAGTTCATCCAGCGCAGTCCCCGCGCCTCCCAGCAGGGTTACTGGGGCTGC 810
          ||| | ||| |||||| | | | ||| |||||||| || | |||
Qy 1122 TCTGTTCCCACCCACTTCAGGGATGCCAGCTTCAGGGCCCCCTTCAGCCATGAC 1181

Db   811 TCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGC 870
          ||||||| || | |||| |||| || || |||||| || ||| || ||||||| ||
Qy 1182 TCCTCCCTACCCGCAGTTTGAGCAATGAACAGGAGAACGGAGACTGTTCATCTGTTTATCCAGC 1241

Db   871 CCAGGGAGTGGGCGCCATCATCGGACAAGAAGGGCAGCACATCAAACAGCTCTCCCGGTT 930
          || | | || || |||||||||| | | ||| |||||| ||||| ||| | || ||
Qy 1242 TCTATCAGTCGGTGCCATCATCGGCACATCGGCAGCACACATCAAGCAGCTTTCTCGCTT 1301

Db   931 TGCCAGGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCGACAAAGTTCGTATGGT 990
          ||| || | | ||||||||||| ||| | ||| ||| | | ||||| ||||||
Qy 1302 TGCTGGAGCTTCAATTAAGATTGCTCCAGCGAAGCACCAGATGCTAAAGTGAGGATGGT 1361

Db   991 TATCATCACTGGACGCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT 1050
          ||||||||||||||| |||||||| ||||||||| ||||| |||||| | |||| |||
Qy 1362 GATTATCACTGGACGACCACCAGAGGCTCAGTTCAGGTCAAGGCTCAGGGAAGAGATTTATGGAAAAT 1421

Db 1051 CAAGGAGGAGAACTCTTTGTCCAAGGAGGAAGTGAAGCTGGAGACCCAATACCGTGT 1110
          ||| ||| |||| | ||| ||| | | || |||||||||| | || |||| ||||
Qy 1422 TAAAGAAGAAAACTTGTTAGTCCTAAAGAAGAGTGAAACTTGAAGCTCATATCAGAGT 1481

Db 1111 GCCAGCATCAGCAGCTGCCCGGGTCATTGGCCAAAGGTGGAAAAACGGTGAACCAGTTGCA 1170
          |||| | || ||||| || | ||||| || ||| |||| |||||||||| ||| ||||
Qy 1482 GCCATCCTTTGCTGCTGCGCAGTGTTATTGGAAGAGGCAAAACGGTGAATGAAGTTTCA 1541

Db 1171 GAATTGACGGCCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCTGATGAGAACGACCA 1230
          |||||| | ||| || ||| | || |||||||||| | ||||||||| ||||
Qy 1542 GAATTTGTCAAGTGCAGAAGTTGTGTTCCTCGTGACCAGACACCTGATGAGAATGACCA 1601

Db 1231 GGTCATCGTGAAAATCATCCGGACCATTTCTATGCCCAGTTCAGATGGCTCAACGAGGATCGG 1290
          ||| | || |||||| || | | | ||||||| |||| ||| | ||| | |||| ||
Qy 1602 AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTGTTGCCCAGTGTGCCCAGGAGAAAAATTCA 1661

Db 1291 AGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAG 1327
          || ||| |||| |||| | ||||||||||| ||| |
Qy 1662 GGAAATTCTGACTCAGGTAAAGCAGCAGCCAACAACAG 1698

RESULT 8
US-09-850-716A-347
; Sequence 347, Application US/09850716A
; Patent No. US20201151339A1
GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERNCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A

; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-850-716A-347

Query Match   34.1%; Score 582.2; DB 9; Length 1740;
Best Local Similarity 66.2%; Pred. No. 6.9e-171;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

Qy   37 AGCCATCATGAAGCTGAATGGCCACCAGTTGAGAACCATGCCCTGAAGGTCTCCTACAT 96
                ||||| |||||| |||||||||||||||| || || |||| ||||||||||| ||||
Db   402 AGCACTAGACAAACTGAATGGATTTCAGTTAGGAGTAGAAGTAGCCTATAT 461

Qy   97 CCCGATGAGCAGATAGACACAGGAGGCTCTGAGAGTGGCGCCCAGGGGGCTTTGGCTCTCG 156
                ||| ||| ||| || |||||||||||||||||||||||||||||||||| |||||||
Db   462 CCCTGATGAAAACGGCCGCCCAGCAAAACCCCTTGCAGCAGCCCCAGCAGCAGCAAGT 521

Qy   157 GGGTCAGCCCCGCCAGGGCTCACCTGTGCAGCGGGGCCCCAGCAGCAGCAGCAAGT 216
                ||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   522 TGGGCAGAGAGGGGTCCTCAAGGCAGGGGTCTCCAGGATCCGTATCCAAGCAGAAACCATG 581

Qy   217 GGACATCCCCTTCGGCTTCTGCGCCCTGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGA 276
                ||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   582 TGATTTGCCTCTGCGCCCTGCTGGTTCCCCAATTTGTTGGAGCCATCATAGGAAAAGA 641

Qy   277 GGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAGATAGAGCGTGCATAGGAA 336
                ||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   642 AGGTGCCACCATTCGGACAACATCACCACAACAGACCCAGTCCACTAAAATCGATGTCCACCGTAA 701

Qy   337 GGAGAACGGCAGGTGCAGCTGAGAAAAGCCATGAGTGCACTCCACCCTGAGGGGCTGCTC 396
                |||||| |||||||||||||||||| |||||| ||||||||||||| |||||||||||
Db   702 AGAAAATCGGGGGCGTGCTGAGAGTGCTGATTACTATCCTCCTCTACTCCTGAGGGCCACCTC 761

Qy   397 CTCCGCTTGTAAGATGAATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACGGC 456
                |||| |||||||| |||||| |||||||||||||||| ||||| |||||||||||||
Db   762 TCCGGCTTGTAAGTCTATTCTGGAGATTATGCATAAGGAGCTCAAGATATAAAATTCAC 821

Qy   457 TGACGAGGTTCCCCTGAAGATCCTGGCCCATATAACTTTGTAGGGCGTCTCATTGGCAA 516
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   822 AGAAGAGATCCCCTTGAAGATTTTAGTCATAATAACTTTGTTGGACGTCTTATTCGGTAA 881

Qy   517 GGAAGGACGGAACCTGAAGAAGGTAGAGCAAGATACCGGACAACAAAATCACCATCTCCTC 576
                ||||||||||||||| ||||||| |||||||| ||||| |||||||||| |||||||
Db   882 AGAAGGAAGAGAAATCTTAAAAAAATTGAGCAAGACACACAGACACTAAAATCACGATATCTCC 941

Qy   577 GTTGCAAGACCCTACCCTTACAACCCTGAGAGCGACCCATCACTGTGAAGGGGGCCATCGA 636
                |||||||||||| ||||||||||||||||||||||||||||| |||||||||||||
Db   942 ATTGCAGGAATTGACGCTGTATAATCCAGAACGCACTAT-ACAGTTAAAGGCAATGTTGA 1001

Qy   637 GAATTGTTGCAGGGGCCCAGCAGAATAATGAAGAAGTTCGGGAGGCCTATGAGAATGA 696
                |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1002 GACATGTGCCAAAGCTGCGAGGGAGGAGATCATGAAGAAGAAAATCAGGGAGTCTTATGAGAAATGA 1061

Qy   697 TGTGGGCTGCCATGA------GCTCTCACCTGATCCCTGGCCTGAACCCTGCTGTAGG 750
                |||||| |||||||| ||||||||||||||||||||| || ||||||||||
Db   1062 TATTGCTTCTATGAATCTTCAAGCACCATTTAATTCCTGGATTAAATCTGAACGCCTTGGG 1121

Qy   751 TCTTTTTCCCAGTTCATCAGCAGCGCCAGTCCCGCGCCTCCCAGCAGCGGTTACTGGGGCTGC 810
                |||||||||||||||| ||||||||||||||||||||||||||||||||| ||||||||
Db   1122 TCTGTTCCCACCCACTTCAGGGATGCCACCTCCCAACCTCAGGGGCCCCTTCAGCCATGAC 1181

Qy   811 TCCCTATAGCTCCTTTATGCAGGCTCTCCGAGCAGGATCGTGCAGGTGCAGGTGTTTATCCCCGC 870
                ||||| |||||||||||||||||||||| ||||| |||||||||||||||||||||||
Db   1182 TCCTCCCTACCCGCAGTTTGAGCAATGCAGGGACACGGACACTGTTCATCTGTTTATCCCAGC 1241

Qy   871 CCAGGCAGTGGGCGGCCATCATCGGCAAGAAGAGGGCAGCACCATCAAACAGCTCTCCCGGTT 930
                |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 TCTATCAGTGGGGCGTGCAAGCAGCAGGGGCCAGCATCAACAATCCGTATCCAAGCAGAAACCATG 1301

Qy   1242 TCTATCAGTGGGGCGTGCAAGCAGCAGGGGCCAGCATCAACAATCCGTATCCAAGCAGAAACCATG 276

; SEQ ID NO 347
; LENGTH: 1740;
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-347

RESULT 9
US-09-897-778-347
; Sequence 347, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-347

Query Match   34.1%; Score 582.2; DB 9; Length 1740;
Best Local Similarity 66.2%; Pred. No. 6.9e-171;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

Qy   37 AGCCATCATGAAGCTGAATGGCCACCAGTTGAGAACCATGCCCTGAAGGTCTCCTACAT 96
                ||||| |||||| |||||||||||||||| || || |||| ||||||||||| ||||
Db   402 AGCACTAGACAAACTGAATGGATTTCAGTTAGGAGTAGAAGTAGCCTATAT 461

Qy   97 CCCGATGAGCAGATAGACACAGGAGGCTCTGAGAGTGGCGCCCAGGGGGCTTTGGCTCTCG 156
                ||| ||| ||| || |||||||||||||||||||||||||||||||||| |||||||
Db   462 CCCTGATGAAAACGGCCGCCCAGCAAAACCCCTTGCAGCAGCCCCAGCAGCAGCAAGT 521

Qy   157 GGGTCAGCCCCGCCAGGGCTCACCTGTGCAGCGGGGCCCCAGCAGCAGCAGCAAGT 216
                ||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   522 TGGGCAGAGAGGGGTCCTCAAGGCAGGGGTCTCCAGGATCCGTATCCAAGCAGAAACCATG 581

Qy   217 GGACATCCTCGGCTTCTGCGCCCTGTGCCCACCCAGTTAAGCAGCATCAGAAG 1327
                |||||||||| |||||||||||||
Db   1662 GGAAATTCTCAGGTAAGCTAAGCAGCACCAACAACAG 1698

Qy   931 TGCCAGCGGCTCCATCAAGAATTGCACCACCCGAAACACCTGACTCCAAAGTCGTATGGT 990
                |||||||| |||||||||||||||| || |||||||||||||||||||||||||||
Db   1302 TGCTGGAGCTTCAATTAAGATTGCTCCAGCGGAGCAACCACCAGATGCTAAAGTGAGGATGGT 1361

Qy   991 TATCATCACTGACCGCCAGAGGCCCAATTCAAGGCTCAGTTCAAGGAGAATCTATGGCAAACT 1050
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 GATTATCACTGACCACCAGAGGCTCAGGCTCAGGAGAAGAATTTATGGAAAAAT 1421

Qy   1051 CAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTGAAGCTGAAGCCCACATACGTGT 1110
                |||||||||||||| ||||||| ||||||||||||||||| ||||||||| |||||||
Db   1422 TAAAGAGAAAAACTTTGTTAGTCCTAAAGACAGAGGTGAAACTTGAAGCTCATCAGAGT 1481

Qy   1111 GCCAGCATCAGCAGCAGGCGCCGGTCATTGGCAGAGTTGGCAAAGTGGAAAAACGGTGAACGAGTTGCA 1170
                |||||| |||||||||||||||| ||||| | |||||||||||| ||||| |||||| |||||
Db   1482 GCCATCCTTTGCTGCTTGGCAGAGTTATTGGAAAGCAAAGCGGTCAAATGAACTTCA 1541

Qy   1171 GAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGACCAGAACCCCTGATGAGAACGACCA 1230
                |||||| |||||| |||| || || |||||||||||||||||||||||| ||||||||
Db   1542 GAATTTGTCAAGTGCAGAAGTGTTGTCCCTCGTGCACCAGAACACCTGATGAGAATGACCA 1601

Qy   1231 GGTCATCGTGAAAAATCATCCGACATTTCTATGCCAGTCAGTCAGATGGCTCAACGGAAGATCCG 1290
                ||| |||| ||||| |||| |||||| ||| | |||||||| || ||||||||||||||
Db   1602 AGTGGTTGTCAAAATAACTGCTCACTTCTATGCTTCATGCTTGCCCAGTTGCCCAGTGAGAAAAATTCA 1661

Db 1662 GGAAATTCTGACTCAGTGAAGCAGCAGCCACCAACAG 1698

RESULT 10
US-10-007-700-347
; Sequence 347, Application US/10007700
; Publication No. US2003006494A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samual X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-007-700-347

Query Match 34.1%; Score 582.2; DB 13; Length 1740;
Best Local Similarity 66.2%; Pred. No. 6.9e-171;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

```
Qy   37 AGCCATCATGAGCGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT   96
          ||| | || ||  | ||||| | ||| ||| || |||| |  |  |||||| | || |
Db  402 AGCACTAGACAAACTGAATGGAGTTTCAGTTAGAGAATTCACCTTGAAAGTAGCCTATAT  461

Qy   97 CCCCGATGAGCAGCAGATAGCACAGGGACCTGAGAATGGGCGCCAGGGGGGCTTTGGCTCTCG  156
          ||| |||||  | ||| | ||  | | ||| ||||| ||| | ||| || | | || || | |
Db  462 CCCTGATGAAACGGCCGCCCAGCAAACCCCTTGCAGCGCGGGGGCCCCAGCAGCAGCGGGGCT  521

Qy  157 GGGTCAGCCCCGCCCAGGGGCTCACCTGTGGCAGCGGGGGCCCAGCCAAGCAGCAGCAAGT  216
          ||| ||| |  | | |||||  | || ||| || | || |  || || ||| |  |  ||
Db  522 TGGGCAGAGGGGGTCCTCAAGGCAGGGGTCTCCAGGATCCGTATCCAAGCAGGAGAAACCATG  581

Qy  217 GGACATCCCCCTTCGGCTTCCGTGTCCTGGTGCCACCCAGTATGTGGGTGCCATTATTGGCAAGGA  276
          |||| ||| | || ||| || || |||  ||| ||||| || | ||| |||||| | |||| |||
Db  582 TGATTTGCCTCTGCGCCTCGGCCCTCGGTTCCCACCCAATTTGTTGGAGCCATCATAGGAAAAGA  641

Qy  277 GGGGGCCACCATCCGCAACATCACAAAACAGACGCCAGTCCAAGATAGACGTGCATAGGAA  336
          | | |||||||| || || ||  | |||| |||| || |  ||| | || | | |||| ||
Db  642 AGGTGCCACCATTCGGACAATCGGACAACATCAACAGACACCGTAAAAATCTAAATCAGAGT  701

Qy  337 GGAGAACGCAGGTGCAGTGCTGAAGTGGAAAAAGCCATCAGTGTGCACTCAGTGTGCA  396
          ||| || | |||  | ||| | || || || || ||||| ||  |   | |   |||
Db  702 AGAAAATGCGGGGGCTGCGTGAGAGTGCATTACTATCCTTCTACTCCTGAAGCACCTC  761

Qy  397 CTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGAGGCTAAGACACCAAAACGGC  456
          | | ||||||||| | || | |||||| | |||| | ||  |||| ||| || || ||||
Db  762 TGCGGCTTGTAAGTCTATTCTGGAGATTATGCATGAGATTATGCATAAAGAAGCTCAAGATATAAAATTCAC  821

Qy  457 TGACGAGGTTCCCCTGAAGATCCTGGCCCATATATAACTTGTAGGGCGTCTCATTGGCAA  516
          |||| ||||| | |||||  | ||||| |||| |  | || ||| | | || |||| ||
Db  822 AGAAGGATCCCCTTGAAGATTTTGACGTCTATATAATAACTTTGTTGGACGTCTTATTGGTAA  881

Qy  517 GGAAGGACGGAACCTGAGAAGCTGAGAGCAAGATAACCGGACAAAATCACCATCTCTC  576
          || | | ||||  || | || | ||||| ||| | | || | ||| |||| ||| | |
Db  882 AGAAGGAAGAAATCTTAAAAAAATTGAGCAAGACACACAATCACGATATCTCC  941

Qy  577 GTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCATCACTGTGAAGGGGGCCATCGA  636
          ||||| || | | |||| || || |||| ||| | ||||| || ||| ||| | | ||| ||
Db  942 ATTGCAGGAATTGACGCTGTATAATCCAGAAACGCCACTATTACAGTTAACAGCAATGTTGA  1001

Qy  637 GAATTGTTGCAGGGCGCGAGCGAGAAATATGCAGAGAAAGTTCGGGAGGCTATGAGAATGA  696
          ||| || || || | | ||| || || | || || || | || | ||| ||| | ||||
Db 1002 GACATGCTGCCAAAGCTGCGGAGGAGAGAGATCATGAGAATCAGGAGAGTCTTATGAAAATGA  1061

Qy  697 TGTGGCTGCCATGA------GCTCTCACCTGATCCCTGGCCTGAACCTTGGCTGCTGTAGG  750
          || ||| |||||||      ||||| | | |||  |   | |||| | ||| | || | ||
Db 1062 TATTGCTTCTATGAATCTTCAAGCACATTTAATTCCTGGATTAAATCTGAACGCCCTTGGG  1121

Qy  751 TCTTTTCCCAGCTTCATCCAGCGATGCCACTCCACTCCCACCTCCAGGGCCCCTTCAGCCATGAC  810
          ||| || | ||| | || ||| | || | |||| | || |||||||||| || | ||||| |
Db 1122 TCTGTTCCACCCACTTCAGGGATGCCACCGCCGCTCCAGGGCCCCCGTTCATCTGTTTATCCGAC  1181

Qy  811 TCCCTATAGCTCCTTTATGCAGGTCCCCGAGCAGGAGATGGTGCAGTGTTTATCCCGC  870
          ||||| | |||| | ||||||  | || ||||| || | ||| |||| | ||||  |||
Db 1182 TCCTCCCTACCCGGCAGTTGAGCAATCAGAAAACGGAGACTTGTTCATCTGTTTATCCAGC  1241

Qy  871 CCAGGCAGTGGCGCCATCATCGGCAAGAGGGGCAGCACAGAGTCTCTCCCGGTT  930
          || || ||| |||| ||| ||||| | || ||||| | || ||  || ||| ||
Db 1242 TCTATCAGTCGTGCCATCATCGGCAGCAGGGGCAGCAGCATCATCAAGCAGCTTTCTCGGTT  1301

Qy  931 TGCCAGCGCCTCCATCAAGATTGCACCACCGAAACACCTGACTTCGTATGGT  990
          ||||| | ||| || ||||| |  | ||| |||  | || ||||| | |||
Db 1302 TGCTGGAGCTTCAATTAAGATTGCTCAGCGGAACCACCAGATGCTAAAGTAGTGACGATGGT  1361

Qy  991 TATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAGGAATCTATGGCAAACT  1050
          || || || |||||| |||| || ||| | |||||| ||||| | | |||| ||| |||
Db 1362 GATTATCACTGGACCGCCAGAGGCTCAGGGAGGAGAATTTATGAAAAAT  1421

Qy 1051 CAAGAGGAGGAGACTTCTTCGTCCAAGGAGGAGAGTCTGAGAGCTGAGACCCACCATACGTGT  1110
          |||||| || || || ||| | ||| |||| || || | || ||| | | |  || | | |
Db 1422 TAAAGAAGAAACTTTGTAGTCCTAAAGAAGAGTGAAACTTGAAGCTCATATCAGAGT  1481

Qy 1111 GCCAGCATCAGCAGCTGGCCGGGTCATTGGCGGGTCATTGGGAAAGTGGAAAACGAGTTGCA  1170
          |||| || ||||| | ||| |||| | |||  |||| | |||||| || || | ||||||||
Db 1482 AGGTGCCACCATTCGGACAACATCACCAAACAACAGACCCGTGAAAAAGCATCAGTGTGCACTCAGTGTGCA  1541

Qy 1171 GAATTTGACGGCAGTGGTAGTAGCCAAGACCCTGATGAGAACGACCA  1230
          ||| ||||  | | || ||||| || || |||| || ||| ||| |
Db 1542 GAATTTGTCAAGTCAAGTGGCAGAAGTGCAGAAGCTCATCATCGGACACGACTTATGGCAAACGTCAGCAGATGAATGAACCA  1601

Qy 1231 GGTCATCGTGAAAATCATCATCGGACATTCTATCTCATGCGTCAACGGAGACCG  1290
          | ||| ||| |||| || |||| || |  | |  |||||   || | ||| ||
Db 1602 AGTGGTTGTCAAAATAACTGGTCACTTCACTTCATGCTTGCCCATGAGAGAAATTCA  1661

Qy 1291 AGACATCCTGCCCAGTTAAGCAGCAGCAGCATCAGAAG  1327
          || || ||| |  | || || |||| | ||  || | |
Db  822 AGAAGAGATCCCCTTTGAACGATTTTAGCTCCATAATAACTTTGTTGGACGTCTCATAATTGGTAA  881
```

```
;  APPLICANT: Wang, Tongcong
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C18
;  CURRENT APPLICATION NUMBER: US/10/117,982
;  CURRENT FILING DATE:  2002-04-05
;  NUMBER OF SEQ ID NOS: 484
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 347
;    LENGTH: 1740
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-117-982-347

Query Match             34.1%;    Score 582.2;  DB 15;  Length 1740;
Best Local Similarity   66.2%;    Pred. No. 6.9e-171;
  Matches  858;  Conservative    0;  Mismatches  433;  Indels    6;  Gaps    1;
```

```
Qy      37 AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT   96
           |||||||||||||||||| |||||||   ||||||||   ||||||||||||  ||||||
Db     402 AGCACTAGACAAACTGAATGATTTCAGTTAGAGAATTTCACCTTGAAAGTAGCCTATAT  461

Qy      97 CCCGATGAGCAGATAGCACAGGGACCTGAGAATGGCGCCGAGGGGCTTTGGCTCTCG   156
           |||||  |||||||||||||  |||  |||||||||||||||||||  |||||||
Db     462 CCCTGATGAAACGGCCGCCCAGCAGAAACCCTTGCAGCGGGGGGCCCGAGGTCGGCGGGGGCT  521

Qy     157 GGGTCAGCCCCGCCAGGGCTCACCTGTGCAGCGGGGGCCCAGCCAGCAGCAGCAGAGT   216
           |||||||  |||  |||||||||  |||||||||    ||||||||||| |||||
Db     522 TGGGCAGAGGGGGCTCCTCAAGGCAGGGGTCTTCCAGATCCGTATCCAAGCAGAAACCATG  581

Qy     217 GGACATCCCCTTCGGCTCGTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCCAAGGA   276
           |||||||||||  ||||  ||||||  |||||||  ||| ||  ||||||||| |||||
Db     582 TGATTTGCCTCTGCCGCCTGGTTCCCACCCAATTTGTTGTGGAGCATCATAGAGAAAAGA  641

Qy     277 GGGGGCCACCATCCGCAACATCACAAAACGAGCCCAGTCCAGTCAAGATAGACGTGCATAGAA   336
           |||||||||||  |||||||  ||||  |||||  |||||  ||||||||  |  |||
Db     642 AGGTGCCACCATTCGGAACATCACCAACAGACCCCAGTCTAAAATCGATGTCCACCGTAA  701

Qy     337 GGAGAACGCAGCAGCTGAAGCTGCTGAGCATGGCCATCAGTTGAGACCCTGAGGGGCTGCTC   396
           ||||||||  |  |||||  ||  |||||  ||||  |||  |||||||  ||||||
Db     702 AGAAAATGCGGGGCGGTGCTGAGAAGGGTGGAAAAACGGTAAGGACACCCAAAACGGC  761

Qy     397 CTCCGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACGGC   456
           |||||  ||||  |||||||  ||||  ||||||  |||||||  |||||||||
Db     762 TGCGGCTTGTAAGTCTATTCTGGAGAATTATGCATAAGGAGCTCAAGATATAAAATTCAC  821

Qy     457 TGACCAGGTTCCCCTGAAGATCCTGGCCCATAAATAACTTGTAGGCGTCTCCATTGGCAA   516
           |||||||||  ||||||||||  |||  ||||  |||||||   ||  ||  |||||
Db     822 AGAAGAGATCCCCTTGAAGATTTTTAGTCTATAATAACTTTGTTGGACGTCTTATTGGTAA  881

Qy     517 GGAAGGACGGAACCTGAACGGTAGAGCAAGATACCGAGAGACAAAAAATCACCATCTCCTC   576
           ||||||||  |||||||||  ||||  ||||||||||||||||||||||||  ||
Db     882 AGAAGGAAGAAATCTTAAAAAATTGAGCAAGACAACTAAAATCACGATATCTCC  941

Qy     577 GTTGCAAGACCTTACCCTTTACAACCTGAAGGGGGCCATCGA   636
           ||||||||||||||  ||||||||||  ||||||||||||
Db     942 ATTGCAGGAATTGACGCTGTATATTCCGAACCGCACCACTATTACAGTTAAAGGCAATGTTGA 1001

Qy     637 GAATTGTTGCAGGGCCGAGCAGCAGGAAATAATGAGCAAAGTTCGGGAGGCCTATGAGAATGA   696
           |||||  |||||  |||  ||||||||||  |||  ||||  |||||||  ||  ||||||
Db    1002 GACATGTGCCAAAGCTGAGGAGGAGATCATGAGAAATCAGGAGAAATCAGGAGTCTTATGAAAATGA 1061

Qy     697 TGTGCGTGCCATGA   750
           ||||  |||||||
Db    1062 TATTGCTTCTATGAATCTTCAAGCACATTTAATTCCTGGATTAAATCTGGGGGCTGC  1121

Qy     751 TCTTTTCCAGCTTCATCCAGCGAGTTTGAGCAATGCAGTCTTCCCCGGT   810
           ||||||||||||  ||  ||||||||  |||||||||    ||||||
Db    1122 TCTATCAGTCGGTCCATCATCGGGATGCCACTCCCACTTCAGGGCCCCCTTCAGCCATGAC 1181

Qy     811 TCCCTATAGCTCCTTATGCAGCAGCATCAGAAG  1327
           ||||||  |||  ||||||||  ||||||
Db    1182 TCCTCTACCCGCAGTTTGAGCAATCAGAAACGGAGACTGTTCATCGTGTTTATCCCAGC 1241
```

```
RESULT 11
US-10-117-982-347
;  Sequence 347, Application US/10117982
;  Publication No. US20030138438A1
;  GENERAL INFORMATION:
;  APPLICANT:  Foy, Teresa M.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Mericle, Barbara
;  APPLICANT:  Spies, Gregory A.
;  APPLICANT:  Fan, Liqun
```

```
Qy    871 CCAGGCAGTGGGGCGGCCATCATCGGCAAGAGGGCAGCACATCAAACAGCTCTCCCGGTT 930
              ||||| ||||||||| |||||||| ||||||
Db   1242 TCTATCAGTCGGTGCCATCATCGGCAAGCAGGGCGGCCAGCATCATCAAGCAGCTTTCTCGCTT 1301

Qy    931 TGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTCGTATGGT 990
              |||||||||||||
Db   1302 TGCTGGAGCTTCAATTAAGATTGCTCCCGGAAGCGAAGCACCAGATGCTAAAGTGAGGATGGT 1361

Qy    991 TATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT 1050
              |||||||| |||||
Db   1362 GATTATCACTGGACCACCACCAGAGGCTCAAGGCTCAGGGAAGAATTTATGGAAAAAT 1421

Qy   1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTTGGAGACCCCACATACGTGT 1110
              ||||| |||||||
Db   1422 TAAAGAAGAAAACTTTGTTAGTCCTAAAGAAGAGGTGAAACTTGAAGCTCATATCAGAGT 1481

Qy   1111 GCCAGCATCAGCAGCTGCGCCGGTCATTGGAACGGTGAACGAGTTGCA 1170
              |||||||
Db   1482 GCCATCCTTGCTGCTGCAGAGTTATTGGAAAGGAGGCAAAACGGTGAATGAACTTCA 1541

Qy   1171 GAATTTGACGGCCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGAACGACCA 1230
              ||||||||| |||||||
Db   1542 GAATTTGCAAGTGCAGAGTTGTTGTCCTCGTCAGTCATTGTTGGACGTCTTATTGTAA 1601

Qy   1231 GGTCATCGTGAAAATCATCGGACATTCTTCATGCCAGTCAGTCAGATGGCTCAACGGAAGATCCG 1290
              |||||||
Db   1602 AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTTGGGCCAGTTGCCCAGTGAATGAAGGAAAATTCA 1661

Qy   1291 AGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAG 1327
              |||||||
Db   1662 GGAAATTCTGACTCAGGTAAAGCAGCACCAACAACAG 1698
```

RESULT 12
US-10-117-982-478
; Sequence 478, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 478
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-982-478

Query Match   34.1%; Score 582.2; DB 15; Length 1740;
Best Local Similarity   66.2%; Pred. No. 6.9e-171;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

```
Qy     37 AGCCATCATGAGCTGAATGGCCACCAGTTGGACGAACCATGCCTTGAGACCGTCTCCTACAT 96
              |||||||||||
Db    402 AGCACTAGACAAACTGAATGAATGGATTTCAGTTAGACGAATTTCAACCTTGAAAGTAGCCTTATAT 461

Qy     97 CCCCGATGAGCAGAGATAGCCACAGGGACCTGAGAATGGGCGCCGAGGAATGGGCTTTGGCTCTCG 156
              |||||||
Db   1542 GAATTTGTCAAGTGCAGAGAAGTTGTGTCCCTCGTGACCAGCACACTGACCA 1601
```

```
Db    462 CCCTGATGATGAAACGGCCCCCAGCAAACCCCCTTGAGCAGCCCCGAGTGGCGCCGGGGGCT 521
Qy    157 GGGTCAGCGCCGCCAGGGCTCACCTGTGGCAGGGCGGCCCAGCAAGCAAGCAGCAGCAAGT 216
Db    522 TGGGCAGAGGGGCTCCTCCAAGGCAGGGGTCTCCAGGATCCGTATCCAAGCAGCAGAAACCATG 581
Qy    217 GGACATCCCCCTCGGCTCTGGTGCTCCACCCAGTATGGGTGCCATTATTGGCAAGGA 276
Db    582 TGATTTGGCTCTGCGGCCTGCTTGCTGCACCCAATTTGTTGAGCCATCATAGGAGAAAGA 641
Qy    277 GGGGGCCACCATCGGCAACATCACAAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAA 336
Db    642 AGGTGCGCCACCATTCGGAACATCACCAACAGACCCAGTCTCTAAAATGTTCCACCGTAA 701
Qy    337 GGAGAACGCAGGTGCAGCTGCTGAAAAGCCATCAGTGTGCACTCACCCCTGAGGGGCTGCTC 396
Db    702 AGAAAATGCGGGGCGTGCTGCAGAGTCGATTACTATCCTCTCACTCCTGAAGGCACCTC 761
Qy    397 CTCCGCTTGTAAGATGATCTTGGAGATATGCATAAAGAGGGTAAGGACACCAAAACGGC 456
Db    762 TGCGGCTTGTAAGTCTATTCTGGAGATTATGCATAAAGAAGGAAGCTCAAGATATAAAAATTCAC 821
Qy    457 TGACGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTGTCATTTGTTGGACGTCTTCATTGGCAA 516
Db    822 AGAAGAGACATCCCCTTGAAGATTTAGCTCATAATAACTTTGTTGGAGTTTATTGTAA 881
Qy    517 GGAAGGACGGAACCTGCAAGGTAGAGCTAGACCAAGAAGATACCGAGACAAAAATCACCATCTCCTC 576
Db    882 AGAAGGAAGAAATCTTAAAAAATTGAGCAAGACCATCTATTACAGATCAATCAGCAACTAAAATCACGATATCTCC 941
Qy    577 GTTGCAAGACCTTACCCTTACAACCTGAGAGACCATCACTGTGAAGGGGCCATCGA 636
Db    942 ATTGCAGGAATTGACGCTGTATAATCCAGAACGAACGCACTATTACAGTTAAAGGCAATGTTGA 1001
Qy    637 GAATTGTTGCAGGAGCCAGCAGCAGGAAAATAATGAAGAGAAGTTCGGGAGGCCTATGAGAATGA 696
Db   1002 GACATGTGCCAAAGTCAGGCGAGATCGAGGAGATCATGAGAGAAAATCACGGGAGTCTTATGAAAAATGA 1061
Qy    697 TGTGTGGCGCCATGA------GCTCTCACCTGATCCCTGCAACCTGGCTGCTGGCCTGCTGGTCTAGG 750
Db   1062 TATTGCTTCTATGAATCTTCAAGCACATTTAATTCCTGGATTAAATCTGGACGCGCCTTGG 1121
Qy    751 TCTTTTCCCAGCTTCATCCAGCGACGGCCAGTCCCCGCCTTCCCCACCTCAGGGTTACTCGGGGCTGC 810
Db   1122 TCTGTTCCCACCTTCAGGGATGCCAGTTCAGGGGAGCCCCTTCAGGGCCCGCGTCAGCCATGAC 1181
Qy    811 TCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAGATGTGCAGGTGTTTATCCCCGC 870
Db   1182 TCCTCCCTACCCGGCAGTTGAGCAATCAGAAACGGAGACTGTTCATCTGTTTATCCCCAGC 1241
Qy    871 CCAGGCAGTGGGGCGCCATCATCGGCGAAGGAGACGGGGGCACATCAAACAGCTCTCCCGGTT 930
Db   1242 TCTATCAGTCGGTGCCATCATCGGGCAAGCAGGGCGGGCCAGCATCATCAAGCAGCAGCTTTCTGCTT 1301
Qy    931 TGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTCGTATGGT 990
Db   1302 TGCTGGAGCTTCAATTAAGATTGCTCCCGGAAGCGAAGCACCAGATGCTAAAGTGAGGATGGT 1361
Qy    991 TATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT 1050
Db   1362 GATTATCACTGGACCACCACCAGAGGCTCAAGGCTCAGGGGAAGAATTTATGGAAAAAT 1421
Qy   1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGGAAGTGGAGACCCCACATACGTGT 1110
Db   1422 TAAAGAAGAAAACTTTGTTAGTCCTAAAGAAGAGGTGAAACTTGAAGCTCATATCAGAGT 1481
Qy   1111 GCCAGCATCAGCAGCTGGCCGGGTCATTGGCCGGGTGAACGTGAACAGGTGAACGAGTTGCA 1170
Db   1482 GCCATCCTTGCTGCTGCAGAGTTATTGGAAAAGGAGGCAAAACGGTGAATGAACTTCA 1541
Qy   1171 GAATTTGACGGCGCAGCTGAGGTGGTAGTACCAAGAGACCAGACACCCCTGATGAGAACGACCA 1230
Db   1542 GAATTTGTCAAGTGCAGAGAAGTTGTGTCCCTCGTGACCAGCACACTGACCA 1601
```

Qy   577 GTTGCAAGACCTTACCCTTTACAACCCTGAGAGACCATCACTGTGAAGGGGGCCATCGA 636
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   942 ATTGCAGGAATTGACGCTCGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGA 1001

Qy   637 GAATTGTTGCAGGGCCGAGCAGGAAATATGAGAGAAGTTCGGAGAAAGTTCGGGAGAATGA 696
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GACATGTGCAAAGTGAGGAGGATCATGAGAGAAAATCAGGGAGTCTTATGAAAATGA 1061

Qy   697 TGTGGCTGCCATGA------GCTCTCACCTGATCCCGGCTGAACCTGGCTGCTGTAGG 750
           ||||||||||||| ||| |||||||||||||||||||||||||||||||||||||
Db   1062 TATTGCTTCTATGAATCTTCAAGCACATTTAATTCCTGGATTAAATCTGAACGGCCTTGG 1121

Qy   751 TCTTTTCCCAGCTTCATCGGCAGTCCGCGCCGCTCCCAGCAGCGTTACTGGGCGTGC 810
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1122 TCTGTTCCCACCACTTCAGGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCCATGAC 1181

Qy   811 TCCCTATAGCTCCTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGC 870
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1182 TCCTCCCTACCCGCAGTTTGAGCAATCAGAAACGAAAACGGAGACTGTTCATCTGTTTATCCCAGC 1241

Qy   871 CCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGCAGCACATCAAACAGCTCTCCCGGTT 930
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1242 TCTATCAGTGGTGCCATCATCGGCAAGCAGCAGGCCGCAGCACATCAAGCACAGCTTTCTCGCTT 1301

Qy   931 TGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAGTTCGTATGGT 990
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1302 TGCTGGAGTTCAATTAAGATTGCTCCACGCAGGACCACCAGATGCTAAAGTGAGGAGGATGT 1361

Qy   991 TATCATCACTGGACCGCCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATGCAAACT 1050
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1362 GATTATCACTGGACCACCACCAGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGAAAAAT 1421

Qy   1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGAGTGAGAGCGTGAAGGGCCACACATACGTGT 1110
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1422 TAAAGAAAGAAAACTTTGTTAGTCCTAAAGAAGGAGTGAAACTTGAAGCTCATATCAGAGT 1481

Qy   1111 GCCAGCATCAGCAGCTGCGCGGGTCATTGGCAAAGGTGAAAAAACGGTGAACGAGTTGCA 1170
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1482 GCCATCCTTTGCTGCTGCGGCAGAGTTATTGGAAAAGAGGAGCAAAACGGTGAATGAACTTCA 1541

Qy   1171 GAATTTGACGCAGCAGCTGAGGTGGTAGTACCAAGAGACACCCTGATGAGAACGACCA 1230
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1542 GAATTTGTCAAGTGCAGAAGTTGTTGTCCTGTGACCAGACACCTGATGAGAATGACCA 1601

Qy   1231 GGTCATCGTGAAAATCATCGGACATTTCATGCCAGTCAGATCAGATGGCTCAACGGAGATCCG 1290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1602 AGTGGTTGTCAAAATAACTGGTCACTTTATGTCTTCCTTTATGCTGCCAGGTTGCCAGGAGAAATTCA 1661

Qy   1291 AGACATCCTGCCCAGGTTAAGCAGCAGCATCAGAAG 1327
           |||||||||||||||||||||||||||||||||||
Db   1662 GGAAATTCTGACTCAGGTAAAGCAGCACCAACAACAG 1698

RESULT 14
US-10-313-986-478
; Sequence 478, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
;   APPLICANT: Foy, Teresa M.
;   APPLICANT: McNabb, Andria
;   APPLICANT: Watanabe, Yoshihiro
;   APPLICANT: Reed, Steven G.
;   APPLICANT: Wang, Tongtong
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C19
;   CURRENT APPLICATION NUMBER: US/10/313,986
;   CURRENT FILING DATE: 2002-12-04
;   NUMBER OF SEQ ID NOS: 560
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 478
;   LENGTH: 1740
;   TYPE: DNA

Qy   1231 GGTCATCGTGAAAATCATCGGACATTTCATGCCAGTCAGATGGCTCAACGGAGGAATCCG 1290
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1602 AGTGGTTGTCAAAATAACTGGTCACTTTGCTCCTATGCTTCTGCCAGGAGCATCAGAAG 1661

Qy   1291 AGACATCCTGGCCCAGGTTAAGCAGCATCAGAAG 1327
           ||||||||||||||||||||||||||||||||
Db   1662 GGAAATTCTGACTCAGGTAAAGCAGCACCAACACAG 1698

RESULT 13
US-10-313-986-347
; Sequence 347, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
;   APPLICANT: Foy, Teresa M.
;   APPLICANT: McNabb, Andria
;   APPLICANT: Watanabe, Yoshihiro
;   APPLICANT: Reed, Steven G.
;   APPLICANT: Wang, Tongtong
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;   TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;   FILE REFERENCE: 210121.455C19
;   CURRENT APPLICATION NUMBER: US/10/313,986
;   CURRENT FILING DATE: 2002-12-04
;   NUMBER OF SEQ ID NOS: 560
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 347
;   LENGTH: 1740
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-313-986-347

  Query Match   34.1%;
  Best Local Similarity   66.2%;
  Matches 858;   Conservative   0;   Mismatches 433;   Indels   6;   Gaps   1;

               Score 582.2; DB 16; Length 1740;
               Pred. No. 6.9e-171;

Qy   37 AGCCATCATGAAGCTGAATGGCCACCAGTTGGAACCATGCCCTGAAGGTCTCCTACAT 96
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   402 AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAATTTCACCTTGAAAGTAGCCTATAT 461

Qy   97 CCCCGATGAGCAGATAGCACGGACACGAGATGGCGCCCAGCAATGGGCGCCTTTGGCTCTCG 156
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   462 CCCTGATGAAACGCGCCGCCAGCAGGTCTCACCTGTGCAGCGGGGCCCCAGCAGCAGCAAGT 521

Qy   157 GGGTCAGCCCCGCCAGGGTCTCCTCAAGGCAGGGGTCTCCAGCAGGCAGGGTCTATCCG 216
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   522 TGGGCAGAGGGGTCCCTTCGGCTGTGGTCACCTGTTCCCACCCAGTATGTGGTGCCATTATTGGCAAGA 581

Qy   217 GGACATCCCCTTCGCCTCTGCGCCTGTGCGCCTCGCCCAATTTGTTGAAGGTAAGCATAGGGAAAAGA 276
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   582 TGATTTGCTCTGCGCCTGTGCGCCTCGCCCAGCAGGGTCTCCAGGGTCTCCAGGATCCGGTGCCATCATG 641

Qy   277 GGGGGCCACCACCATCCGCAACATCCGCAACAGCACAGACAGACCCAGTCCAAGATGACGTGCATAGGAA 336
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   642 AGGTGCCACCATTCGGAACATCCGGAACATCACCAAGACACCCAGTCTAAAATCGATGTGCCACCGTAA 701

Qy   337 GGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCTGAGGGGTGCTC 396
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   702 AGAAAAATGCGGGGGCGTGCTGCTGAGAGTCGATTTCATTCCTTCTTCTCTGAAGGCACCTC 761

Qy   397 CTCCGCTTGTAAGATGATCTTGGAGATTATGCATCAAGAGAGGCTAAGGACACCCAAAACGGC 456
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   762 TCGGCGTTGTAAGTCATTCTGGAGATTATGCATAAGCATAAGGAAGCTCAGATATAAAATTCAC 821

Qy   457 TGACGAGTTCCCCGAAGATCCGCCCATAATAACTTTGTAGGGGGTTCATTGGCAA 516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   822 AGAAGAGATCCCCTTGAAGATTTTAGCTCATAATAACTTTGTTGGACGTCTTATTGGTAA 881

Qy   517 GGAAGGACGGAACCTGAAGAAGGTAGAGCAAGAGCACAAGCACCCAAAATCACCATCCTCC 576
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   882 AGAAGGAGAAGAAAATTGAGCAAGAAGCACAAGCAAGCACAGACACTAAAATCACGATATCTCC 941

```
;. ORGANISM: Homo sapiens
US-10-313-986-478

Query Match       34.1%;   Score 582.2; DB 16; Length 1740;
Best Local Similarity  66.2%;   Pred. No. 6.9e-171;
Matches   858;   Conservative   0;   Mismatches 433;   Indels   6;   Gaps   1;

Qy    37  AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT   96
          ||| |||    |  |||||||  |  ||||| |||| |   ||| |||| || |  | ||
Db   402  AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAATTCACCTTGAAAGTAGCCTATAT  461

Qy    97  CCCCGATGAGCAGATAGCAGAGGACCTGAGGATGGGCGCCGAGCAAAAACCCTTGGCTCTCG  156
          ||    ||||| ||  ||  || || | ||| | |||  |||  ||   |  || ||| |||
Db   462  CCCTGATGAAAACGGCCGCGCCAGCAAACCCCTTGCAGCGGGGCCCCAGCGCAGCAGCGGCT  521

Qy   157  GGGTCAGCCCCCGCCAGGGCTCACCTGTGGCACGCGGGGCCCCAGCCAGCAGCAGCAAGT  216
          || |||   |||   | ||| | |||||||| |||  |||  ||||| ||| |||| |||
Db   522  TGGGCAGGGGGGGTCCTCAAGGCAGGGGTCTCCAGCAGCCGTATCCAAGCAGAAAACCATG  581

Qy   217  GGACATCCCCCTTCGGCTCCTGGTGCCACCCAGTGTGGGTGCCATTATTGGCAAGGA     276
          | ||||| |   | ||  || ||||||| ||||| |||| ||  ||   | || |
Db   582  TGATTTGCCCTCTGCGGCTGGTTCCCACCCAATTGTTGTTGGAGCCATCATAGGAAAAGA  641

Qy   277  GGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAA  336
          || || || |||   |||| | | |||| | |   |||| | ||  | |    |  |
Db   642  AGGTGCCACCATTCGGAACATCACCAAACAGAGATCTAAAATCTAAAATCGATGTCCACCGTAA  701

Qy   337  GGAGAAGGCAGTGCAGCTGAAAAGGCCATCAGTGTGCACTCCACCCCTGAGGGGTGTCTC  396
          || ||  |||| |  | |||||  ||||  |   | |||| || |  |||  |   |   |
Db   702  AGAAAATCGCGGGGGCGTGCTGAGAGTCGATTCAGTATCCTCTTACTACCTCCTGAGGCACCTC  761

Qy   397  CTCCGGTTGTAAGATGATCTTGGAGATTATGCATAAAGACGCTAAGGACACCCAAAAAACGGC  456
          | ||  | |  | | |  || ||||  | |  ||| | ||| ||| || || ||||| | ||
Db   762  TGCGGCTTGTAAGTCTATTCTGAGATTATGCATAAGGAAGCTCAAGATATAAAAATTCAC  821

Qy   457  TGACGAGGTTCCCCTGAAGATCCTGCCCATAATAACTTGTAGGGCGTCTCATTGGCAA    516
          ||  ||||| |||||||||  |  || |  | | | || | || || || ||||  ||
Db   822  AGAAGAGATCCCCCTGAAGATTTTAGTCTCATATAACTTGTTGGACGTCTTATTGTGTAA  881

Qy   517  GGAAGGACGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTC  576
          || |||  ||||||| || ||| | ||| ||| ||| |||| |||||  ||||  | |
Db   882  AGAAGAAGAAATCTTAAAAAATTGAGCAAGAACACAGACACTAAAAATCACGATATCTCC  941

Qy   577  GTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCACTCACGTGAGGGGGCCATCGA   636
          | || ||| || |  || | | | | | || ||  | |  |||  | |  || ||   ||
Db   942  ATTGCAGAATTGACGCTGTATATAATCCAGAACAACAGTTATTACAGTTAAAGGCAATGTTGA  1001

Qy   637  GAATTGTTGCAGGGCCCGAGCAGCAGGAAATAATGAAGAAAGTTCGGGAGGCCATGAGAATGA  696
          || ||| |||| ||| | |||||||| |||  ||||||| | |||  |   ||| || |
Db  1002  GACATGTGCCAAAGCTGAGGAGGAGATCATGAGGAGAAATCAGGAGGTCTTATGAAAATGA  1061

Qy   697  TGTGGCTGCCATGA------GCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAGG  750
          | || ||||||||       ||| || |   | ||  || || |  || || ||| |||
Db  1062  TATTGCTTCTATGAAATCTTCAAGACAATTTAATTCCTGGATTTAAATCTGAACGCCTTGGG  1121

Qy   751  TCTTTTCCCAGCTTCATCCAGCGCAGTCGCCGCCGCCTTACTGGGGCTGC  810
          || | | |||||| | |  ||| || |  | |  ||| | |    ||||
Db  1122  TCTGTTCCCACCCACTTCCAGGGATGCCATCTCCAGGCCATGACTCAGCCATGAC  1181

Qy   811  TCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGC  870
          || ||||  || |  || || | | ||||| || | | | || |  | || | | ||| ||
Db  1182  TCCTCCCTACCCCAGTTTGAGCAATCAGAAACGGAGACTGTTCATCTGTTTATCCCCAGC  1241

Qy   871  CCAGGCAGTGGGGCGCCATCATCCGCAAGAAGGGGCAGCACATCAAAACAGCTCTCCCGGTT  930
          ||||| | ||  |  ||||   |  |||| || |  ||| || |  |||| |   ||  | ||
Db  1242  TCTATCAGTCGGTGCCATCATCGGCAGGGCCAGCCAGCACATCAAGCCAGAGAAACCATG  1301

Qy   931  TGCCAGCCGCCTCCATCAAGATTGCACCACCCGAAAACCTGACTCCCAAAGTTCGTATGGT  990
          || | || |  || || | ||||  |  ||||||||| ||| |||||||| | |    | ||
Db  1302  TGCTGGAGCTTCAATTAAGATTGTGCTCCAGCCGGAAGCCATCATAGAGAAGAAAGA  1361

Qy   991  TATCATCACTGACCGCCAGAGGCCAGATTCCAAGGCTTCAGGAGAAGAATCTATGGCCAAACT  1050
          || || ||  ||| | || ||  | |||  ||| | | ||||  || | | | |  || |
Db  1362  GGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAA  1421
```

```
Db  1362  GGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAA  1421

Qy  1051  CAAGGAGGAGAGAACTTCTTTGGTCCCCAAGGACGGAGGAGAGTGAAGTCGAGACCACCGTGT  1110
          .||||| ||||| ||||||  |||  || || ||| |   |||  | || ||| | ||||
Db  1422  TAAAGAGAAAACTTTGTTAGTCCTAAAGAGAGAGAGGTGAAACTTGAAGCTTGAGATCAGAGT  1481

Qy  1111  GCCAGCATCAGCAGCTCGCCGGGTCATTGGCCAAAGGTGCAAAACGGTGAAAACCGTGAACGAGTGCA  1170
          ||| || ||||| |  |||  ||| |||  | ||||| ||||| || ||| ||| ||||||  |
Db  1482  GCCATCCTTGCTGCTGGCAGAGTTATTGGAAAAGAGGCCAAAAAGGTGAATGAACTTCA  1541

Qy  1171  GAATTGACGGCAGCTGAGGTGGTAGTACCAAGAGACACCCCTGATGAGAACGACCA  1230
          |||||| |  || | ||| || || ||  | ||||||| | ||  | || | ||||
Db  1542  GAATTTGTCAAGTGCAGAAGTTGTGTCCCTCGTGACCAGCACCTGATGACGAATGACCA  1601

Qy  1231  GGTCATCGTGAAAAATCATCGGACACATTTCTATGCCAGTCAGATGCCTCAGTGTGCAAGT  1290
          ||| || | ||| |||| |  || | ||||| | ||||| |  ||  ||| ||| |  ||
Db  1602  AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTGTGCCAGGTTGCCCAGAGAAAAATTCA  1661

Qy  1291  AGACATCCTGGCCCAGGTTAAGCAGCAGCAGCATCAGAAG  1327
          ||| |  |||| | ||| |||||||||| || ||| |||
Db  1662  GGAAATTCTGACTCAGTAAAGCAGCAGCACCAACACAG  1698


RESULT 15
US-09-897-778-447
; Sequence 447, Application US/09897778
; Patent No. US2002147143A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT: Marnerakis, Margarita
;  APPLICANT: Fanger, Gary R.
;  APPLICANT: Vedvick, Thomas S.
;  APPLICANT: Carter, Darrick
;  APPLICANT: Watanabe, Yoshihiro
;  APPLICANT: Henderson, Robert A.
;  APPLICANT: Peckham, David W.
;  APPLICANT: Fanger, Neil
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C16
;  CURRENT APPLICATION NUMBER: US/09/897,778
;  CURRENT FILING DATE: 2001-06-28
;  NUMBER OF SEQ ID NOS: 467
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 447
;  LENGTH: 1743
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-09-897-778-447

Query Match       34.1%;   Score 582.2; DB 9; Length 1743;
Best Local Similarity  66.2%;   Pred. No. 6.9e-171;
Matches   858;   Conservative   0;   Mismatches 433;   Indels   6;   Gaps   1;

Qy    37  AGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT   96
          ||| |||    |  |||||||  |  ||||| |||| |   ||| |||| || |  | ||
Db   402  AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAATTCACCTTGAAAGTAGCCTATAT  461

Qy    97  CCCCGATGAGCAGATAGCAGAGGACCTGAGGATGGGCGCCGAGCAAAAACCCTTGGCTCTCG  156
          ||    ||||| ||  ||  || || | ||| | |||  |||  ||   |  || ||| |||
Db   462  CCCTGATGAAAACGGCCGCGCCAGCAAACCCCTTGCAGCGGGGCCCCAGCGCAGCAGCGGCT  521

Qy   157  GGGTCAGCCCCCGCCAGGGCTCACCTGTGGCACGCGGGGCCCCAGCCAGCAGCAGCAAGT  216
          || |||   |||   | ||| | |||||||| |||  |||  ||||| ||| |||| |||
Db   522  TGGGCAGGGGGGGTCCTCAAGGCAGGGGTCTCCAGCAGCCGTATCCAAGCAGAAAACCATG  581

Qy   217  GGACATCCCCCTTCGGCTCCTGGTGCCACCCAGTGTGGGTGCCATTATTGGCAAGGA     276
          | ||||| |   | ||  || ||||||| ||||| |||| ||  ||   | || |
Db   582  TGATTTGCCCTCTGCGGCTGGTTCCCACCCAATTGTTGTTGGAGCCATCATAGGAAAAGA  641

Qy   277  GGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAA  336
          || || || |||   |||| | | |||| | |   |||| | ||  | |    |  |
Db   642  AGGTGCCACCATTCGGAACATCACCAAACAGAGATCTAAAATCTAAAATCGATGTCCACCGTAA  701
```

Wed Jul 14 08:56:06 2004

us-09-270-437d-5_1.rnpb

```
Qy  337  GGAGAACGCAGGTGCAGCTCGAGCTGAAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTC  396
              ||||| ||||||||||||| ||||||||| ||||| |||| |||||||  |||||||||||||||||
Db  702  AGAAATGCGGGGCTGCTGAGGAAGTCGATTACTATCCTCCTACTCCTGAAGGCACCTC  761

Qy  397  CTCCGCTTGTAAGATGATCTTGGAGATTATGCAGATTATCATAAAGAGCTAAGGACACCAAAACGGC  456
              ||||||||||||||||||||||||||| |||| ||||||| |||||||||||||||||||||
Db  762  TGCGGCTTGTAAGTCTGTTAGTCTATTCTGGAGATTATGCATAAGGAAGCTCAAGATATAAAATTCAC  821

Qy  457  TGACGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAA  516
              |||  ||||||||| ||||||||||||||||||||||||| |||||||| ||||||||
Db  822  AGAAGAGATCCCCCTTGAAGATTTTAGTCATCATAATAACTTTGTTGGACGTCTTATTGGTAA  881

Qy  517  GGAAGGACGGAACCTGAAGAAGGTAGACGAAGCAAGATACCGAGACAAAAATCACCATCTCCTC  576
              ||||| |||| |||||||||| ||||| ||||||||||| |||| ||||| ||| |||||
Db  882  AGAAGAAGAGAAATCTAAAAAAATTGAGCAAGACACAGACACTAAAATCACGATATCTCC  941

Qy  577  GTTGCAAGACCTTACCCTTACAACCCTGAGAGACCATCACTGTGAAGGGGCCATCGA  636
              ||||||||||  |||||||| ||||||||  || |||| ||||||||| |||| |||
Db  942  ATTGCAGGAATTGACCGTGTATAATCCTGATAATCCAGATACAGTTAAAGGCAATGTTGA  1001

Qy  637  GAATTGTTGCAGGGGCCGAGCAGGAGAAATAATGACAGAAAGTTCGGGAGGCCTATGAGAATGA  696
              ||||| ||||||||||||||||||||||||||||  |||| || |||||||||||||| |||||
Db  1002 GACATGTGCCAAAGCTGAGGAGGAGGAGATCATGAAGAGATCATGAAGAAATCAGGGAGTCTTATGAAAAATGA  1061

Qy  697  TGTGGGCTGCCATGA-----GCTCTCACCTGATCCCTGGCCTGGCTGCTGTAGG  750
              |||||  ||||||        ||||| ||| |||  |||| |||  |||| |||
Db  1062 TATTGCTTCTATGAATCTTCAAGCACATCATTAATTCCTGGATTAAATCTGAACGCCTTGGG  1121

Qy  751  TCTTTTCCAGCTTCATCCAGCTTCATCCAGCGCAGTCCCCGCCTCCCACCTCAGGGCCGTTACTGGGGCTGC  810
              ||||| |||||| |||||| |||||| ||||  ||||||||||||| ||||| |||||| |||
Db  1122 TCTGTTCCCACCCACTTCAGGGATGCCACCTCACCCTCAGGGCCCCCTTCAGCCATGAC  1181

Qy  811  TCCCTATAGCTTCCTTTATGCAGGCTCCCGAGCAGGAGATGTGCAGGTGTTTATCCCCGC  870
              ||||| ||| |||| ||||| ||||||| |||||||| |||| || ||||||||||||
Db  1182 TCCTCCCCTACCCGGCAGTTTGAGCAATCAGAAACGGACTGTTCATCTGTTTATCCCCAGC  1241

Qy  871  CCAGGCAGTGGGGCCCATCATCGCAAGAAGGGCAGCACATCAAACAGCTCTCCCGGTT  930
              |||||||| ||| ||||| |||| ||| |||| ||||||||  |||||||||||
Db  1242 TCTATCAGTCGGTGCCATCATCGGCAAGCAGGGCCGGCAGCCTTTCTCGCTT  1301

Qy  931  TGCCAGGCGCCTCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTCGTATGGT  990
              |||||| ||||||||| |||| |||| |||| |||||||| || |||||||||| |||||
Db  1302 TGCTGGACTTCAATTAAGATTGCTCTCAGCGGAAGCGCACCAGATGCTAAAGTGAGGATGGT  1361

Qy  991  TATCATCACTGACACCGGCAGAGGCCCCAATTCAGGGTCAGGGAAGAATCTATGGCAAACT  1050
              |||||||||||| ||||||||||| ||| |||||| |||||||| ||||| ||||||||
Db  1362 GATTATCACTGACGACACCAGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGGAAAAAT  1421

Qy  1051 CAAGGAGGAGAACTTCTTTGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGT  1110
              ||||||| ||| ||||| ||| |||| ||||||||||||||| |||||||||| ||||
Db  1422 TAAAGAGAAAACTTTGTTAGTCCTAAAGAGAGAGGTGAAACTTGAGCTCATATCAGAGT  1481

Qy  1111 GCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGTGGAAAAAACGGTGAACGAGTTGCA  1170
              |||||| ||||||||||| |||||||||| ||||||| ||| || |||||||||||||
Db  1482 GCCATCCTTTGCTGCTGGCCAGAGTTATTGGCAGAGTTTATTGGAAAAAGGAGGCAAAAACGGTGAATGACTTCA  1541

Qy  1171 GAATTTGACGGCGAGCTGAGGTGGTGGTAGTACCAAGAGAACCAGACCCCTGATGAGAACGACCA  1230
              ||||| ||||||| |||| ||||| ||||||| |||  ||||| |||||| |||||||||| |||
Db  1542 GAATTTGTCAAGTGCAGAAGTGTTGTTGTTCCCTCGTCGTACCTGACGACACCTGATGAGAATGACCA  1601

Qy  1231 GGTCATCGTGAAAATCATCGGACACATTCTATGCCAGTCAGATGGCTCAACGGAAGATCCG  1290
              |||||||| |||| |||||| |||| |||| ||| ||||| |||||||| || ||||||
Db  1602 AGTGGTTGTCAAAATAACTGGTCACTTCATATGTTGCCAGTTGCCCAGAGAAAATTCA  1661

Qy  1291 AGACATCCTGGCCCAGGTTAAGCAGCATCAGAG  1327
              ||||||| ||||| |||| |||||  || |||
Db  1662 GGAAATTCTGACTCAGGTAAGCAGCACCAACAACAG  1698
```

Search completed: July 13, 2004, 23:41:03
Job time : 798 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:　July 13, 2004, 14:50:27 ; Search time 699 Seconds
　　　　　　　　　　　　　　　　　(without alignments)
　　　　　　　　　　　　　　　　10380.433 Million cell updates/sec

Title:　　　　　US-09-270-437D-5
Perfect score:　1708
Sequence:　　　1 aggagctgcgcaccgcc........atttcctcaggtttaaaa 1708

Scoring table:　IDENTITY_NUC
　　　　　　　　GapOp 10.0 , Gapext 1.0

Searched:　　　3373863 seqs, 2112099041 residues

Total number of hits satisfying chosen parameters:　6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
　　　　　　　　　Maximum Match 100%
　　　　　　　　　Listing first 45 summaries

Database :　　N_Geneseq_29Jan04:*
　　　　　　　1: geneseqn1980s:*
　　　　　　　2: geneseqn1990s:*
　　　　　　　3: geneseqn2000s:*
　　　　　　　4: geneseqn2001as:*
　　　　　　　5: geneseqn2000bs:*
　　　　　　　6: geneseqn2002s:*
　　　　　　　7: geneseqn2003as:*
　　　　　　　8: geneseqn2000bs:*
　　　　　　　9: geneseqn2003cs:*
　　　　　　　10: geneseqn20004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1703	99.7	1703	3	AAZ36151	Aaz36151 DNA encod
2	1669.2	97.7	1946	3	AAZ36153	Aaz36153 An altern
3	1389	81.3	2780	4	AAK94782	Aak94782 Human ful
4	1161.8	68.0	2224	2	AAZ10617	Aaz10617 cDNA enco
5	1085.4	63.5	1182	4	AAS26148	Aas26148 Human cDN
6	1085.4	63.5	1182	7	ABX73489	Abx73489 Human nov
7	1069.8	62.6	1129	4	AAS26566	Aas26566 Human cDN
8	1069.8	62.6	1129	7	ABX73907	Abx73907 Human nov
9	732.8	42.9	833	4	AAK91969	Aak91969 Human cDN
10	732.8	42.9	833	8	AAK93655	Aak93655 Human cDN
11	582.2	34.1	1740	3	AAC66035	Aac66035 Human lun
12	582.2	34.1	1740	6	ABL49254	Abl49254 Human lun
13	582.2	34.1	1740	6	ABQ92440	Abq92440 Human lun
14	582.2	34.1	1743	6	ABL49299	Abl49299 Human lun
15	582.2	34.1	1743	6	ABL49297	Abl49297 Human lun
16	582.2	34.1	1743	6	ABQ92483	Abq92483 Human lun
17	582.2	34.1	1743	6	ABQ92485	Abq92485 Human lun
18	582.2	34.1	1743	8	ADA28540	Ada28540 Recombina
19	582.2	34.1	1743	8	ADA28537	Ada28537 Recombina
20	582.2	34.1	1764	6	ABL49283	Abl49283 Human lun
21	582.2	34.1	1764	6	ABQ92469	Abq92469 Human lun
22	582.2	34.1	1764	8	ADA28518	Ada28518 Recombina
23	580.6	34.0	1740	8	ADA28437	Ada28437 Human lun
24	580.6	34.0	4159	3	AA236150	Aa236150 DNA encod
25	580.6	34.0	4181	3	AAC65900	Aac65900 Human lun
26	580.6	34.0	4181	6	ABL65405	Abl65405 Lung canc
27	580.6	34.0	4181	6	ABL49119	Abl49119 Human lun
28	580.6	34.0	4181	6	ABQ92305	Abg92305 Human lun
29	580.6	34.0	4181	8	ADA28265	Ada28265 Human lun
30	580.6	34.0	4181	9	ADD14661	Add14661 Human src
31	580.6	34.0	4181	8	ADE53470	Ade53470 Human lun
32	580.6	34.0	4434	8	ACH03940	Ach03940 Human cDN
33	580.6	34.0	4601	6	ABA99958	Aba99958 Human KOC
34	551.4	32.3	1985	5	AAS76779	Aas76779 DNA encod
35	530	31.0	4264	5	AAS86150	Aas86150 DNA encod
36	504	29.5	3283	3	AAZ36154	Aaz36154 An altern
37	504	29.5	3667	6	AAS70982	Aas70982 DNA encod
38	504	29.5	3667	6	ABS76442	Abs76442 cDNA enco
39	489.2	28.6	2010	5	AAS70981	Aas70981 DNA encod
40	489.2	28.6	3412	5	AAZ36152	Aaz36152 DNA encod
41	489.2	28.6	3694	4	AAS26150	Aas26150 Human cDN
42	489.2	28.6	3694	7	ABX73491	Abx73491 Human nov
43	454	26.6	1707	7	ACA90176	Aca90176 cDNA enco
44	267.2	15.6	2290	4	AAS26153	Aas26153 Human cDN
45	267.2	15.6	2290	7	ABX73494	Abx73494 Human nov

ALIGNMENTS

RESULT 1
AAZ36151
ID　AAZ36151 standard; DNA; 1708 BP.
XX
AC　AAZ36151;
XX
DT　11-FEB-2000　(first entry)
XX
DE　DNA encoding cancer associated antigen KOC-2.
XX
KW　Cancer associated antigen; KOC-2; cancer; vaccine; CT7; ss.
XX
OS　Homo sapiens.
XX
PN　WO9954738-A1.
XX
PD　28-OCT-1999.
XX
PP　16-MAR-1999;　99WO-US005766.
XX
PR　17-APR-1998;　98US-00061709.
XX
PA　(LUDW-) LUDWIG INST CANCER RES.
XX
PI　Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
XX
DR　WPI; 2000-013284/01.
XX
PT　Nucleotides representing cancer-associated genes, used to develop
PT　products for the diagnosis, monitoring and treatment of cancers.
XX
PS　Claim 55; Page 40; 44pp; English.
XX
CC　The present sequence represents a cancer associated antigen gene
CC　designated KOC-2. The specification also describes a cancer associated
CC　antigen designated CT7. The CT7 polynucleotide was isolated from SK-MEL-
CC　37 melanoma cells. The polypeptide has some homology with MAGE-10,
CC　limited to about 210 carboxy terminal amino acids. The amino terminal of
CC　the protein has a repetitive pattern, with repeats rich in serine,
CC　proline, glutamine and leucine, and an almost invariable core of the
CC　peptide given in AAY43877. The CT7 polypeptide can be processed to
CC　peptides which provoke lysis by cytolytic T cells. The polynucleotides
CC　and polypeptides can be used for treating a cancerous condition and
CC　screening for or diagnosing cancerous conditions. The cancer associated
CC　antigens can be used as an immunogenic or vaccine composition with an
CC　adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony

```
CC  stimulating factor (GM-CSF)
XX
SQ  Sequence 1708 BP; 447 A; 469 C; 473 G; 314 T; 0 U; 5 Other;

Query Match      99.7%;  Score 1703;  DB 3;  Length 1708;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1708;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AGGGACGCTGCCGCCACCGCCCCAGTTACCCCGGGAGCCATCATGAAGCTGAATGCCA     60
          ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db      1 AGGGACGCTGCCGCCACCGCCCCAGTTTACCCCGGGAGCCATCATGAAGCTGAATGCCA     60

Qy     61 CCAGTTGGAGAACCATGCCCTGAAGGTCTTCTACATCCCGATGAGCAGATAGCACACAGGG    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACATCCCGATGAGCAGATAGCACACAGGG    120

Qy    121 ACCTGAGAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTCAGCCCCGCCAGGGGCTCACC    180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACCTGAGAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTCAGCCCCGCCAGGGGCTCACC    180

Qy    181 TGTGGCAGCGGGGGCCCCAGCCAAGCAGTGACAGCAGCATCCGCTCGGCTCCTGGT    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TGTGGCAGCGGGGGCCCCAGCCAAGCAGTGACAGCAGCATCCGCTCGGCTCCTGGT    240

Qy    241 GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGAGGGGCCACCATCCGCAACATCAC    300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGAGGGGCCACCATCCGCAACATCAC    300

Qy    301 AAAACAGACCCAGTCCAAGATAGACGTGCATAGGGAAGGAGGACGCAGTGCAGCTGAAAA    360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAAACAGACCCAGTCCAAGATAGACGTGCATAGGGAAGGAGGACGCAGTGCAGCTGAAAA    360

Qy    361 AGCCATCAGTGCACTCACCCCTGAGGGCTGCTCCTCCGCTTGTAAGATGATCTTGGA    420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGCCATCAGTGCACTCACCCCTGAGGGCTGCTCCTCCGCTTGTAAGATGATCTTGGA    420

Qy    421 GATTATGCATAAAAGAGGCTAAGGACACCAAAAACGGCGTGACGAGGTTCCCTGAGATCCT    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GATTATGCATAAAAGAGGCTAAGGACACCAAAAACGGCGTGACGAGGTTCCCTGAGATCCT    480

Qy    481 GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAGGT    540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAGGT    540

Qy    541 AGAGCAAGATACCGAGACAAAAATCACCATCTCCTCTGTTGCAAGAACCTTACCCTTTACAA    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGAGCAAGATACCGAGACAAAAATCACCATCTCCTCTGTTGCAAGAACCTTACCCTTTACAA    600

Qy    601 CCCTGAGGAGGACCATCACTGTGAAGGGCCATCGAGAATTGTTGCAGGGGCCGAGCAGGA    660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCCTGAGGAGGACCATCACTGTGAAGGGCCATCGAGAATTGTTGCAGGGGCCGAGCAGGA    660

Qy    661 AATAATGAAGAAAGTTCGGAGAGAAAGTTCGGAGAGATGATGTGGCTGCCATGAGCTCTCACCT    720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AATAATGAAGAAAGTTCGGAGAGAAAGTTCGGAGAGATGATGTGGCTGCCATGAGCTCTCACCT    720

Qy    721 GATCCCTGCCTGAACCTGGCTGCGTGTAGGTCTTTTCCCAGCTTCATCCAGCGCCAGTCCC    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GATCCCTGCCTGAACCTGGCTGCGTGTAGGTCTTTTCCCAGCTTCATCCAGCGCCAGTCCC    780

Qy    781 GCCGCCTCCCAGCAGCCGTTACTGGGCGCTGCTCCTATAGCTCCTTTATGCAGCTCCCGA    840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GCCGCCTCCCAGCAGCCGTTACTGGGCGCTGCTCCTATAGCTCCTTTATGCAGCTCCCGA    840

Qy    841 GCAGGAGATGGTGCAGGTGTTTATCCCCGCCAGGCAGTGGGCGCCATCATCGGCCAAGAA    900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GCAGGAGATGGTGCAGGTGTTTATCCCCGCCAGGCAGTGGGCGCCATCATCGGCCAAGAA    900

Qy    901 GGGGCAGCAGCATCAAACAGCTCTCCCGGTTTGCCAGCCGCCCTCCATCAAGATTGCACCACC    960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GGGGCAGCAGCATCAAACAGCTCTCCCGGTTTGCCAGCCGCCCTCCATCAAGATTGCACCACC    960

Qy    961 CGAAACACCTGACTCGACTCCAAAGTTCGTATGGTTATCATCACTGGACACCGGAGCGCCCAATT   1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGAAACACCTGACTCGACTCCAAAGTTCGTATGGTTATCATCACTGGACACCGGAGCGCCCAATT   1020

Qy   1021 CAAGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGA   1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAAGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGA   1080

Qy   1081 GGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCGCGGTCATTGG   1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCGCGGTCATTGG   1140

Qy   1141 CAAAGGTGGAAAAACCGGTGAACGAGTTGACGGCAGTGGAGTGGTAGTACC   1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAAAGGTGGAAAAACCGGTGAACGAGTTGACGGCAGTGGAGTGGTAGTACC   1200

Qy   1201 AAGAGACCAGACCCCTGATGAGAACACCAGGTCATCGTGAAAATCATCGGACATTTCTA   1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AAGAGACCAGACCCCTGATGAGAACACCAGGTCATCGTGAAAATCATCGGACATTTCTA   1260

Qy   1261 TGCCAGTCAGATGGCTCAACGAGATCCGAGACATCCTGCCCAGGTTAAGCAGCAGCA   1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGCCAGTCAGATGGCTCAACGAGATCCGAGACATCCTGCCCAGGTTAAGCAGCAGCA   1320

Qy   1321 TCAGAGGGACAGAGTAACCAGGCCCAGGGAAGTGACCAGGGAGTGACCAGCCCTCCTGTCC   1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 TCAGAGGGACAGAGTAACCAGGCCCAGGGAAGTGACCAGGGAGTGACCAGCCCTCCTGTCC   1380

Qy   1381 CTTNGAGTCCAGGACAACAACAACGGCAGAAATCGAGAGTTGCTCTCCCCGCAGGCCTGA   1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CTTNGAGTCCAGGACAACAACAACGGCAGAAATCGAGAGTTGCTCTCCCCGCAGGCCTGA   1440

Qy   1441 GAATGAGTGGGAATCCGGACACNTGGCCGGGCTGTAGATCAGGTTTGCCCACTTGATT   1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GAATGAGTGGGAATCCGGACACNTGGCCGGGCTGTAGATCAGGTTTGCCCACTTGATT   1500

Qy   1501 GAGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCACCCAATTGGC   1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GAGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCACCCAATTGGC   1560

Qy   1561 CCAACACTGNTGCCCTCGGGGTGTCAGAAAATNTAGCGCAAGGAGGGTGATCACCAGTGGGAAGA   1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCAACACTGNTGCCCTCGGGGTGTCAGAAAATNTAGCGCAAGGAGGGTGATCACCAGTGGGAAGA   1620

Qy   1621 ATTGTTTAAAGAAGCTCTCAGCCCCACCAAGAGGGTGATCACACCTCAGTGGGAAGA   1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 ATTGTTTAAAGAAGCTCTCAGCCCCACCAAGAGGGTGATCACACCTCAGTGGGAAGA   1680

Qy   1681 AAAATAAAATTTCCTTCAGTTTTAAAA   1708
          |||||||||||||||||||||||||||
Db   1681 AAAATAAAATTTCCTTCAGTTTTAAAA   1708

RESULT 2
AAZ36153
ID   AAZ36153 standard; DNA; 1946 BP.
XX
AC   AAZ36153;
XX
DT   11-FEB-2000   (first entry)
XX
DE   An alternative form of DNA encoding cancer associated antigen KOC-2.
XX
KW   Cancer associated antigen; KOC-2; cancer; vaccine; CT7; ss.
XX
OS   Homo sapiens.
XX
PN   WO9954738-A1.
XX
PD   28-OCT-1999.
XX
PF   16-MAR-1999;   99WO-US005766.
XX
PR   17-APR-1998;   98US-00061709.
```

XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
XX
DR   WPI; 2000-013284/01.
XX
PT   Nucleotides representing cancer-associated genes, used to develop
PT   products for the diagnosis, monitoring and treatment of cancers.
XX
PS   Claim 55; Page 42; 44pp; English.
XX
CC   The present sequence represents an alternative form of a cancer
CC   associated antigen gene designated KOC-2. The specification also
CC   describes a cancer associated antigen designated CT7. The CT7
CC   polynucleotide was isolated from SK-MEL-37 melanoma cells. The
CC   polypeptide has some homology with MAGE-10, limited to about 210 carboxy
CC   terminal amino acids. The amino terminal of the protein has a repetitive
CC   pattern, with repeats rich in serine, proline, glutamine and leucine, and
CC   an almost invariable core of the peptide given in AAY43877. The CT7
CC   polypeptide can be processed to peptides which provoke lysis by cytolytic
CC   T cells. The polynucleotides and polypeptides can be used for treating a
CC   cancerous condition and screening for or diagnosing cancerous conditions.
CC   The cancer associated antigens can be used as an immunogenic or vaccine
CC   composition with an adjuvant, e.g. a cytokine, a saponin, or granulocyte
CC   macrophage-colony stimulating factor (GM-CSF)
XX
SQ   Sequence 1946 BP; 502 A; 528 C; 553 G; 358 T; 0 U; 5 Other;

Query Match   97.7%; Score 1669.2; DB 3; Length 1946;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   30 CCCGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT 89
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT 327

Qy   90 CCTACATCCCGATGAGCAGCAGATAGCACAGGACCTGAGATGGGCGCCGAGGGGGCTTTG 149
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   328 CCTACATCCCGATGAGCAGCAGATAGCACAGGACCTGAGATGGGCGCCGAGGGGGCTTTG 387

Qy   150 GCTCTCGGGGTCAGCCCCGCCAGGGCTCACCTGTGCCAGGCGCGGGGGCCCAGCCAAGCAGC 209
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   388 GCTCTCGGGGTCAGCCCCGCCAGGGCTCACCTGTGCCAGGCGCGGGGGCCCAGCCAAGCAGC 447

Qy   210 AGCAAGTGACATCCCCCTTCGGCTCTGGCTGTGCCACCCAGTATGTGGGTGCCATTATTG 269
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   448 AGCAAGTGACATCCCCCTTCGGCTCTGGCTGTGCCACCCAGTATGTGGGTGCCATTATTG 507

Qy   270 GCAAGGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGC 329
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   508 GCAAGGAGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGC 567

Qy   330 ATAGGAAGGAGAACGCAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGG 389
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   568 ATAGGAAGGAGAACCCAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGG 627

Qy   390 GCTGCTCCTCGCGTTGTAAGACATGTCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 449
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   628 GCTGCTCCTCGCGTTGTAAGACCTTTACCCTTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 687

Qy   450 AAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCATAATAACTTTGTAGGGCGTCTCA 509
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   688 AAACGGCTGACGAGGTTCCCTGAAGATCCTGGCCATAATAACTTTGTAGGGCGTCTCA 747

Qy   510 TTGGCAAGGAACGGAACCTGAAGAAGGTAGAGCAAGCAAGATACCGAGACAAAAATCACCA 569
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   748 TTGGCAAGGAACGGAACCTGAAGAAGGTAGAGCAAGCAAGATACCGAGACAAAAATCACCA 807

Qy   570 TCTCCTGGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGG 629
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   808 TCTCCTGGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGG 867

Qy   630 CCATCGAGAATTGTTGCAGGGCCCAGCAGCAGGAAATAATGAAGAAGTTCGGGAGCCTATG 689
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   868 CCATCGAGAATTGTTGCAGGGCCCAGCAGCAGGAAATAATGAAGAAGTTCGGGAGCCTATG 927

Qy   690 AGAATGATGTGGCTGCCATGAGCTCTCACCTGAGCTCCTGGCTGCTGCCTGAACCTGCTGTAG 749
         |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   928 AGAATGATGTGGCGCTGCCATGAGCTCTCACCTGAGCTCCTGGCTGCTGCCTGAACCTGCTGTGCTGTAG 987

Qy   750 GTCTTTTCCCAGCTTCATCCAGCGCAGTCCCCGCCGGCCCCGCCGCAGCGTTACGGGGGTG 809
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   988 GTCTTTTCCCAGCTTCATCCAGCGCAGTCCCCGCCGGCCCCGCCGCAGCGTTACTGGGGCTG 1047

Qy   810 CTCCCTATAGCTCCTTTATGCAGCGTCCTCCGACAGGAGGATGGTGCAGTGTGTTATCCCG 869
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1048 CTCCCTATAGCTCCTTTATGCAGCGTCCTCCGACAGGAGGATGGTGCAGTGTGTTATCCCG 1107

Qy   870 CCCAGGCAGTGGGCGCCATCATCCGGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGT 929
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1108 CCCAGGCAGTGGGCGCCATCATCCGGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGT 1167

Qy   930 TTGCAGGCCCTCCATCAAGATGCACCACCCGAAAACACCTGACTCCAAAGTTCGTATGG 989
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1168 TTGCCAGGCCCTCCATCAAGATTGCACCACCCGAAAACACCTGACTCCAAAGTTCGTATGG 1227

Qy   990 TTATCATCACTGACCGCCAGAGGCCCAATTCAAGGTCAGGGAAGAAGATCTATGGCAAAC 1049
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1228 TTATCATCACTGACCGCCAGAGCCGCCCAATTCAAGGCTCAGGGAAGAAGATCATGGCAAAC 1287

Qy 1050 TCAAGGAGGAGAACTTCTTGGTCCCGGGTCATTGGCAAAGGTGGAAAAACGGTGAACGAGTTGC 1109
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1288 TCAAGGAGGAGAACTTCTTTGGTCCCGGGTCATTGGCAAAGGTGGAAAAACGGTGAACGAGTTGC 1347

Qy 1110 TGCCAGCATCAGCAGCTGCGCCGGTCATTGGCGGGGTCGCCGGGGTTGTAGTACCAGGAACCAGACC 1169
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1348 TGCCAGCATCAGCAGCTGCGCCGGTCATTGGCGGGGTCGCCGGGGTGGTAGTACCAGGAACCAGACC 1407

Qy 1170 AGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCAGACCCCTGATGAGAACGACC 1229
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1408 AGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCAGACCCCTGATGAGAACGACC 1467

Qy 1230 AGGTCATGTGAAAATCATGTGGACATTCTATGCCAGTCAGTGGCTCAACGGAAGATCC 1289
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1468 AGGTCATGTGAAAATCATGTGGACATTCTATGCCAGTCAGTGGCTCAACGGAAGATCC 1527

Qy 1290 GAGACATCCTGCCCAGTTAAGCAGCATCAGAGGGACATCAGAGGGACAGAGTAACCAGGCCCAGG 1349
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1528 GAGACATCCTGCCCAGTTAAGCAGCATCAGCAGCATCAGAGGGACAGCAGAGTAACCAGGCCCAGG 1587

Qy 1350 CACGGAGGAAGTGACCAGCCCCTTCCCTGTCCCTTNGAGTCCAGGACAACAACAGGGCAGAA 1409
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1588 CACGGAGGAAGTGACCAGCCCCTTCCCTGTCCCTTNGAGTCCAGGACAACAACAGGGCAGAA 1647

Qy 1410 ATCGAGAGTGTGCTGCTCCCCGGAGGCCTGAGAATGAGTGGGAATCCGGGACANTGGGC 1469
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1648 ATCGAGAGTGTGCTGCTCCCCGGAGGCCTGAGAATGAGTGGGAATCCGGGACANTGGGC 1707

Qy 1470 CGGGCTGTAGATCAGGTTGCCCACTTGATTGCAGAAAGATGTTCCAGTGAGGAGACCCTGA 1529
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1708 CGGGCTGTAGATCAGGTTTGCCCACTTGATTGCAGAAAGATGTTCCAGTGAGGAGACCCTGA 1767

Qy 1530 TCTNTCAGCCCAAACACCCAACACTGTNGCCCTCGGGGTGTCAG 1589
         ||||||||||||||||||||||||||||||||||||||||||||
Db 1768 TCTNTCAGCCCAAAACACCCAACACTGTNGCCCTCGGGGTGTCAG 1827

Qy 1590 AAATTNTAGGCGAAGGCACTTTTAAACCTGAATTGTTTAAACGGAAGAGCTCTCCAGGCCCCAC 1649
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1828 AAATTNTAGGCGAAGGCACTTTTAAACCTGAATTGTTTAAAGAGAGCTGAAGAGCTCTCCAGGCCCCAC 1887

Qy 1650 CAAGAGGGTGGATCACACCTCAGTGGGAAGAAAATAAAATTTCCTTCAGGTTTTAAAA 1708
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1888 CAAGAGGGTGGATCACACCTCAGTGGGAAGAAAATAAAATTTCCTTCAGGTTTTAAAA 1946

RESULT 3
AAK94782
ID   AAK94782   standard; cDNA; 2780 BP.
XX
AC   AAK94782;
XX
DT   06-NOV-2001 (first entry)
XX
DE   Human full-length cDNA, SEQ ID NO: 3886.
XX
KW   Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1130094-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-00114089.
XX
PR   08-JUL-1999; 99JP-00194486.
PR   11-JAN-2000; 2000JP-00118774.
PR   02-MAY-2000; 2000JP-00183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI   Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR   WPI; 2001-524255/58.
DR   P-PSDB; AAM93826.
XX
PT   830 Primers useful for synthesizing full length cDNA clones and their use
PT   in genetic manipulation.
XX
PS   Claim 8; SEQ ID NO 3886; 1380pp + Sequence Listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been isolated
CC   and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC   been determined. Primers for synthesising the full length cDNA are useful
CC   for clarifying the function of the protein encoded by the cDNA. The full
CC   length clones were obtained by construction of full length enriched cDNA
CC   libraries that were synthesised by the oligo-capping method. The primers
CC   enable the production of the full length cDNA easily without any special
CC   methods. The present sequence is a full length human cDNA of the
CC   invention. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in CD-ROM format directly
CC   from EPO
XX
SQ   Sequence 2780 BP; 768 A; 681 C; 671 G; 660 T; 0 U; 0 Other;

Query Match   81.3%;   Score 1389;   DB 4;   Length 2780;
Best Local Similarity 99.2%;   Pred. No. 0;
Matches 1405;   Conservative 0;   Mismatches 5;   Indels 6;   Gaps 1;

Qy   299 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAGGAAGGAGACGCAGGTGCAGCTGAA 358
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAGGAAGGAGACGCAGGTGCAGCTGAA 60

Qy   359 AAAGCCATCAGTGTGCACTCCACCTCCACCCCCTGAGGGCTGCTCCTGTCTGTAAGATGATCTTG 418
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAAGCCATCAGTGTGCACTCCACCTCCACCCCCTGAGGGCTGCTCCTGTCTGTAAGATGATCTTG 120

Qy   419 GAGATTATGCATAAAGAGAGGCTAAGGACCACAAAACGGCTAAAACGGCTGACGAGGTTCCCCTGAGATC 478
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAGATTATGCATAAAGAGAGGCTAAGGACCACAAAACGGCTAAAACGGCTGACGAGGTTCCCCTGAGATC 180

Qy   479 CTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGACCATCCTCCTCGGGGTGTCAGAAATTNAGCGAAGG 538
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGACCATCCTCCTCGGGGTGTCAGAAATTCAGCGAAGG 240

Qy   539 GTAGAGCAAGATACCGAGACAAAAATCACCATCTCCGTTGCAAGACCTTACCCTTTAC 598
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   241 GTAGAGCAAGATACCGAGACAAAAATCACCATCTCCGTTGCAAGACCTTACCCTTTAC 300

Qy   599 AACCCTGAGAGGACCATCACTGCAGAATTGTTGCAGGGCCGAGCAG 658
          |||||||||||||||||||||||||||||--------|||||||||
Db   301 AACCCTGAGAGGACCATCACTGTGCAGAATTGTTGCAGGGCCGAGCAG 360

Qy   659 GAAATAATGAAGAAAGTTCGGGAGGCCATGAGAATGAGATGTGGCTGCCATGAGC------ 712
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GAAATAATGAAGAAAGTTCGGGAGGCCATGAGAATGAGATGTGGCTGCCATGAGCTGCTGCAG 420

Qy   713 TCTCACCTGATCCCTGGCCTGCGTCGTGTAGGTCTTTCCCAGCTTCATCCAGC 772
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCTCACCTGATCCCTGGCCTGCGTCGTGTAGGTCTTTCCCAGCTTCATCCAGC 480

Qy   773 GCAGTCCGCCGCCCTCCCAGCAGCGCTTACTGGGCGTGCTCCTATAGCTCCTTTATGCAG 832
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GCAGTCCGCCGCCCTCCCAGCAGCGCTTACTGGGCGTGCTCCTATAGCTCCTTTATGCAG 540

Qy   833 GCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCAGGCAGTGGGCGCCATCATC 892
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCAGGCAGTGGGCGCCATCATC 600

Qy   893 GGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCCGGTTTGCCAGCGCCTCCATCAAGATT 952
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCCGGTTTGCCAGCGCCTCCATCAAGATT 660

Qy   953 GCACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCGCCAGAG 1012
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GCACCACCCGAAACACCTGACTCCAAAGTTCATGTTATCATCACTGGACCGCCAGAG 720

Qy   1013 GCCCAATTCAAGGCTCAGGGGAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGT 1072
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GCCCAATTCAAGGCTCAGGGGAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGT 780

Qy   1073 CCCAAGGAGAAGTGAAGCTGGAGACCCACATAGTGTGCCAGATCAGCAGTGGCCGG 1132
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 CCCAAGGAGAAGTGAAGCTGGAGACCCACATAGTGTGCCAGATCAGCAGTGGCCGG 840

Qy   1133 GTCATTGGCAAAGTGGAAAAACGTGAACGAGTTGCAGAAATTGACGGCAGTGAGGTG 1192
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GTCATTGGCAAAGTGGAAAAACGTGAACGAGTTGCAGAAATTGACGGCAGTGAGGTG 900

Qy   1193 GTAGTACCAAGAGACCAGACCCCTGATGAGAACGACCAGTCATCGTGAAAATCATCGGA 1252
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GTAGTACCAAGAGACCAGACCCCTGATGAGAACGACCAGTCATCGTGAAAATCATCGGA 960

Qy   1253 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCCTGCCCAGTTAAG 1312
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCCTGCCCAGTTAAG 1020

Qy   1313 CAGCAGCATCGAAGGACAGAGGACGTAACCAGGCACGGAGGAAGTGACCAGCCCCT 1372
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CAGCAGCATCGAAGGACAGAGGACGTAACCAGGCACGGAGGAAGTGACCAGCCCCT 1080

Qy   1373 CCCTGTCCCTTNGAGTCGAGTCCAGGACAACAACGGGCAGAAATCGAGAGTGTGCTCTCCCCGGC 1432
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCCTGTCCCTTCGAGTCGAGTCCAGGACAACAACGGGCAGAAATCGAGAGTGTGCTCTCCCCGGC 1140

Qy   1433 AGGCCTGAGAATGAGTGGGAATCCGGGACACNTGGGCCGTGTAGATCAGGTTTGCCC 1492
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AGGCCTGAGAATGAGTGGGAATCCGGGACACCTGGGCCGTGTAGATCAGGTTTGCCC 1200

Qy   1493 ACTTGATTGAGAAAGATGTTCAGTGAGGACCTGATCNTCAGCCCCAAACACCACC 1552
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACTTGATTGAGAAAGATGTTCAGTGAGGACCTGATCCTCAGCCCCAAACACCACC 1260

Qy   1553 CAATTGGCCCAACACTGTNTGCCCTCGGGGTGTCAGAAATTNATGCGAAGGCACTTTT 1612
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CAATTGGCCCAACACTGTCTGCCCTCGGGGTGTCAGAAATTCATGCGAAGGCACTTTT 1320

Qy   1613 AAACGTGGATTGTTTAAAGAGAGCTCTCAGGCCTCCACCCAAGAGGGTGGATCACACCTCAG 1672
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

us-09-270-437d-5_1.rng

```
Db   1321 AAACGTGGATTGTTTAAGAAGCTCTCCAGGCCCACCAAGAGGGTGGATCACACTCAG 1380

Qy   1673 TGGGAAGAAAAATAAAATTTCCTTCAGGTTTTAAAA 1708
               ||||||||||||||||||||||||||||||||||||
Db   1381 TGGGAAGAAAAATAAAATTTCCTTCAGGTTTTAAAA 1416
```

RESULT 4
AAZ10617
ID   AAZ10617 standard; cDNA; 2224 BP.
XX
AC   AAZ10617;
XX
DT   17-NOV-1999 (first entry)
XX
DE   cDNA encoding a murine c-myc coding region determinant binding protein.
XX
KW   c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;
KW   endonucleolytic attack; half-life; breast cancer; colon cancer;
KW   pancreatic cancer; ss.
XX
OS   Mus musculus.
XX
FH   Key   Location/Qualifiers
FT   CDS   131..1864
FT   /*tag= a
XX
PN   WO9945594-A2.
XX
PD   16-SEP-1999.
XX
PF   05-MAR-1999; 99WO-US004897.
XX
PR   09-MAR-1998; 98US-0077372P.
XX
PA   (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ross J;
XX
DR   WPI; 1999-551506/46.
DR   P-PSDB; AAY30649.
XX
PT   Diagnosing presence or absence of a tumor in a human by examining c-myc
PT   coding region determinant-binding protein.
XX
PS   Example; Fig 1A-D; 79pp; English.
XX
CC   The present sequence encodes a murine c-myc coding region determinant
CC   binding protein (CRD-BP). The presence or absence of a tumor can be
CC   determined by determining the levels of CRD-BP present in the suspect
CC   tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack
CC   and so prolongs its half-life. The methods are used for diagnosing
CC   presence or absence of a tumor in a human, especially breast, colon and
CC   pancreatic cancer. They are also used to inhibit cancer cell growth
XX
SQ   Sequence 2224 BP; 580 A; 609 C; 618 G; 417 T; 0 U; 0 Other;

Query Match   68.0%; Score 1161.8; DB 2; Length 2224;
Best Local Similarity 82.7%; Pred. No. 0;
Matches 1409; Conservative 0; Mismatches 262; Indels 33; Gaps 6;

Qy   30 CCCGGGGAGCCATCATGAAGCTGAAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCT 89
          || | ||||||| ||||||||||||||||||| |||||||||||||||| |||||||||
Db   525 CCAGGCAAGCTATCATGAAGAAGCTAAATGGCCATCAACTGGAGAACCATGCCCTGAAGGTCT 584

Qy   90 CCTACATCCCGATGAGCAGCAGATAGCACAGCAGGGACCTGAGACATGGGCGCCGAAATGGG 149
          |||||||||||||||||||| |||||||| |||||||||||||||||||||||||||||||
Db   585 CCTACATACCTGATGAGCAGCAGATAACAACAAGGTCCTGAGAATGGGCGTCGTGGTGAGGCCTTTG 644

Qy   150 GCTCTCGGGGTCAGCCCCGCCACAGGGTCCCCCGGTCCCCGTGGGCGGCAGCAGGGCTCCAGCAAGTCT 209
          |||||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||||||
Db   645 GGTCTCGGGCGCCAGCCCCGCCAGGGTCCCCCGGTCCCCGTGGGCGGCAGCAGGGGCTCCAGCAAGCAGC 704
```

<table>
<tr><td>Db</td><td>210</td><td>AGCAAGTGGACATCCCCCTTCGGCTCCTCTGGTGCTGGCCACCCAGTATGTGGGTGCCATTATTG</td><td>269</td></tr>
<tr><td>Qy</td><td>705</td><td>AGCCAGTGGACATCCCTCCTCCGGCTCCTCTGGTGCTCCTGGTGCCTACGCAGTATGTGAGGCGCTATCATTG</td><td>764</td></tr>
</table>

```
Qy    270 GCAAGGAGGGGCCACCATCCGCAACATCACAAAGCAGACCAGTCCAGTAGAGACGTGC 329
              |||||||||||||||||||||||||||||||||||||||||||||  |||||||
Db    765 GCAAGGAGGAGGTGCCACCATCCGCAAACATCACAAAACAGACCAGCCAGTCCAAAATAGAGACGTGC 824

Qy    330 ATAGGAAGGAGAACGCAGGTGCAGTCAGTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGG 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    825 ATAGGAAGGAGAGAATGCCGGCGCGCGTGCCGGAGAAGGCCATCAGCGTGCATTCAACCCCTGAAG 884

Qy    390 GCTGCTCCGGCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCA 449
          ||  ||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    885 GCTGCTTCCTCCGGCGTGCAAGATGATCTTGGAGATTATGCACAAGGAGGCAAGGACCACCA 944

Qy    450 AAACGGCTGACCAGAGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCTCA 509
          ||   |||||||||||||||||||||||||||||||||||||| |||||||||||  |||
Db    945 AAACGGCAGATGAAGTTCCCCTGGAAGATCCTGGCCTATAACAACTTCGTCGGGACTCA 1004

Qy    510 TTGGCAAGGAGGACGGAACCTGAAGAAGGTAGAGCAAGTACCGGACAAGAAAATCACCA 569
          |||||||||||||||| |||||||||||||||||||||||||||||||  |||||||||
Db   1005 TTGGCAAGGAGGCCGGAACCTGAAGAAGGTGGAGCAGGACCACAGAGACGAAGATCACCA 1064

Qy    570 TCTCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGG 629
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1065 TCTCATCGCTCCAGGACCTTCACGCTCCTATAACCCTGAGAGGACCATCACTGTGAAGGGCG 1124

Qy    630 CCATCAGAATTGTTGCAGGGCCGAGCAGCAGAAATATGAAGAAATATTCGGGAGGCCTATG 689
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1125 CCATTGAGAACTGTTGCAGGGCCGGAGCAGCAGCCATCATGAAGAAGTTCGAGAGGGCTTACG 1184

Qy    690 AGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCTGCCTGAACCTGGCTG 743
          |||||||||||||||||||||||      |||||||||||||||||||||||||||||
Db   1185 AGAACGACGTGGCCGCCATGAGCTTGCAGTCCATCCCTCATCCCTGGCTTAACCTGCTG 1244

Qy    744 CTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCTCCCCAGCAGCGTTACTG 803
          ||  |||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db   1245 CTTGAGGTCTTCCCCAGCTTCATCCAGCGCAGTCCCAGCGTGTCCTCCTCCCCAGCAGTGTCACTG 1304

Qy    804 GGGCTGCTCCCTATAGCTCCTTATGCAGGCTCCCGAGCAGCAGGATGGTGCAGGTGTTTA 863
          |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||
Db   1305 GGGCTGCTCCCTATAGCTCCTTCCTTCATGCAGCGTCCGGAGCAGGAGATGTACAAGTGTTCA 1364

Qy    864 TCCCCGCCAGCAGTGGCGCCCATCATCGGCCAAGAGAGGGCAGCACATCAAACAGCTCT 923
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   1365 TCCCCGGCCAGCTGTGTGGGGCGCCCATCATTGGCAAGAGGGCCATCAAACAACTCT 1424

Qy    924 CCCGGTTGCCAGCGCCTCATCAAGATTGCCACCACCCGAAACACCTGACTCCAAAGTTC 983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1425 CCCGTTTCGCCAGCGCCTCCATCAAGATTGCTCCACCAGAAACACCTGACTCCAAAGTTC 1484

Qy    984 GTATGGTTATCATCATCGTGACCGCCAGAGCCCAATTCAAGGCTCAGGGAGGAAGAATCTATG 1043
          | || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1485 GAATGGTCGTCATCATCGTGACCGCCAGAGCCCAGTTCAAGGCTCAGGGAGGAAGAATTTATG 1544

Qy   1044 GCAAACTCAAGAGAGAGAACTTCTTTGTCCAAGGAGGAGTGAGCTGAGACCCCACA 1103
          |||||||||||||||||||||  ||||||||||||||||||||  ||||||||||
Db   1545 GCAAACTAAAAGAAGAGAATTTCTTTGTCCCAAGGAGGAGTAAAGCTAGAGACCCCACA 1604

Qy   1104 TACGTGTGCCAGCATCAGCACTGGCCGCGGTCATTGCAAGGTGAAAAACGTGAACG 1163
          ||||||||||||||||||||||||||||||||||| ||||||||||  |||||||
Db   1605 TACGGGTTCCGGCTTCAGCAGCCGCCCGTCATCGGCAAAGGCGGAAAAACGGGGAATG 1664

Qy   1164 AGTTGCAGAATTTGACGGCAGCAGCTGAGGTGGTAGTACCAAGAGACCCTGATGAGA 1223
          ||||||||||||| ||||||||||| |||||||| |||||||||||||||||||||
Db   1665 AGCTGCAGAACTTGACTGCGCTGAGGTGGTAGTGCCAAGAGACCACCCCGGATGAGA 1724

Qy   1224 ACGACCAGGTCATCGTGAAAAATCATCGGACATTTCATGCCAGTCAGATGGCTCAACGGA 1283
          |||||||||||||||||||||||| ||||  |||||||||||||||||||||||||||
Db   1725 ACGACCAAGTCATTGTTAAGATCATCGGACATTTCATGCCAGCCAGCCAGCCAAGCAGC 1784

Qy   1284 AGATCCGAGACATCCTGGCCCAGTTAAGCAGCAGCATCAGAAGGGACGAAGTAACCAGG 1343
```

```
Db        |||||||||||||||| ||| |||||||| |||||| |||| |||||||||
     1785 AGATCCGAGACATCCTGGCTCAGTTAAGCAACAGCACCAGGAGAGCAACCTGG 1844

Qy   1344 CCCAGGCACGGAGGAAGTGA-CCAGCCCCTCCCGTCCCTTNGAGTCCAGGACAACAACG 1402
          |||||||||||||||||| | |||||||||||||||||| |||||||| ||||||||||
Db   1845 CCCAGGCACGGAGGAAGTGACCCCCCTCCTGTCCCATTGGCTCCAAGATCAGCAGG 1904

Qy   1403 GGCAGAA----------------ATCGAGAGTGTGCTCTCCCCGGCCAGGCCTGAG 1441
          |||||||                |||||||||||||||||||||||||||||||||
Db   1905 AGGAACACAGAACTGGAGGGGCGGGGTGGAGGGCGTGTGTTTTTCCCAGCCAGGCCTGAG 1964

Qy   1442 AATGAGTGGGAATCCGGGACACNTGGGCTGGGCTGTAGATCCAGGTTTGCCCACTTGATTG 1501
          ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   1965 AATCAGTGGGAATCAGGG-CATTTGGGCTCGGCTGGGAGATCAGGTTTGCACACTGTATTG 2023

Qy   1502 AGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCACCCAATTGGCC 1561
          |||||||||||||||||||||||||| ||| ||||||||||||||||||||||||||||
Db   2024 AGAACAATGTTCCAGTGAGGAGGAATCCTGATCCTCTCGCCCCCAA--TTGAGCCAGCTGGCCA 2081

Qy   1562 CAACACTGTNTGCCCCTGGGGTGTCAGAAATTNTAGCGCAAGGCACTTTTAAACGTGGA 1621
          ||||||||| | |||||||||||||||||||| ||| |||||||||| |||||||||||
Db   2082 CAGCCCACCCCTTGGAATATCACCATTGCAATCATAGCTTCGCCCCAA TTGGGTTGCTTTTAAACGTGGA 2141

Qy   1622 TTGTTTAAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTGGGAAGAA 1681
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||  |||||||
Db   2142 TTGTCT--TGAAGTTCTCCAGCCTCCATGGAAGGAAGGGTCAGATCCCAGTGGGGAAGAG 2199

Qy   1682 AAATAAAATTTCCTTCAGGTTTTA 1705
          ||||||||||||||||||||||||
Db   2200 AAATAAAAATTTCCTTCAGGTTTTA 2223
```

RESULT 5
AAS26148
ID AAS26148 standard; cDNA; 1182 BP.
XX
AC AAS26148;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 327.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmalogical; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX

```
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   14-AUG-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0245519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0252270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0257758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0271182P.
PR   23-AUG-2000; 2000US-0277009P.
PR   30-AUG-2000; 2000US-0288924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0214413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0233397P.
PR   14-SEP-2000; 2000US-0233398P.
PR   14-SEP-2000; 2000US-0233399P.
PR   14-SEP-2000; 2000US-0234400P.
PR   14-SEP-2000; 2000US-0234401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0230064P.
PR   14-SEP-2000; 2000US-0230065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0249997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
```

PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   05-DEC-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251330P.
PR   05-DEC-2000; 2000US-0251998P.
PR   05-DEC-2000; 2000US-0256719P.
PR   05-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251855P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-488783/53.
DR   P-PSDB; AAU16161.
XX
PT   New nucleic acid molecules encoding 461 human secreted proteins for
PT   diagnosing, preventing, treating or ameliorating medical conditions and
PT   used as food additives or preservatives.
XX
PS   Claim 1; SEQ ID NO 327; 980pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules and their
CC   encoded secreted proteins. The nucleic acids and proteins are used to
CC   prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC   rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC   in diagnosing a pathological condition or susceptibility to a
CC   pathological condition. Antibodies to the proteins can also be used in
CC   alleviating symptoms associated with the disorders and in diagnostic
CC   immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC   (ELISA). Disorders which are diagnosed or treated include autoimmune

CC   diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC   neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC   arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC   nervous system disorders e.g. Alzheimer's disease, infections caused by
CC   bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC   and many other disorders listed in the specification. The polypeptides
CC   can also be used to aid wound healing and epithelial cell proliferation,
CC   to prevent skin aging due to sunburn, to maintain organs before
CC   transplantation, for supporting cell culture of primary tissues, to
CC   regenerate tissues and in chemotaxis. The polypeptides can also be used
CC   as a food additive or preservative to increase or decrease storage
CC   capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC   minerals, cofactors and other nutritional components. The present
CC   sequence encodes a novel secreted protein of the invention. Note: The
CC   sequence data for this patent did not form part of the printed

Query Match   63.5%; Score 1085.4; DB 4; Length 1182;
Best Local Similarity   98.9%; Pred. No. 3.9e-297;
Matches 1102; Mismatches   0; Conservative   0; Indels   6; Gaps   1;

Qy   601 CCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   18 CGCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGA 77

Qy   661 AATAATGAAGAAAGTTCGGGAGGCCTATAGAGAATGATGTGCTGCCATGAGC-----TC 714
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   78 AATAATGAAGAAAGTTCGGGAGGCCTATAGAGAATGATGTGCTGCCATGAGCCTGCAGTC 137

Qy   715 TCACCTGATCCCTGGCCTGAACCTGCTGCGTAGGTCTTTTCCCAGCTTCATCCAGCGC 774
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   138 TCACCTGATCCCTGGCCTGAACCTGCTGCGTAGGTCTTTTCCCAGCTTCATCCAGCGC 197

Qy   775 AGTCCCGCCGCCTCCCAGCAGCGGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGC 834
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198 AGTCCCGCCGCCTCCCAGCAGCGGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGC 257

Qy   835 TCCCGAGCAGGAGATGTGTCAGGTGCAGGTGCAGGTGTTTATCCCCCGGTTGCCAGCGCCT 894
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258 TCCCGAGCAGGAGATGTGTCAGGTGCAGGTGCAGGTGTTTATCCCCCGGTTGCCAGCGCCT 317

Qy   895 CAAGAAGGGGCAGCACATCAAACAGCTCTCCGGTTGCCAGCCGCCTCCATCAAGATTGC 954
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318 CAAGAAGGGGCAGCACATCAAACAGCTCTCCGGTTGCCAGCCGCCTCCATCAAGATTGC 377

Qy   955 ACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACGCCAGAGGC 1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   378 ACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACGCCAGAGGC 437

Qy   1015 CCAATTCAAGGCTCAGGGAGGAATCTATGGCAAAACTCATGGCAAACTCAAGGAGAGACTTCTTTGGTCC 1074
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438 CCAATTCAAGGCTCAGGGAGGAATCTATGGCAAAACTCATGGCAAACTCAAGGAGAGACTTCTTTGGTCC 497

Qy   1075 CAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGTGGCGCGGGT 1134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   498 CAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGTGGCGCGGGT 557

Qy   1135 CATTGGCAAAGTGGAAAAACGGTGAAACCGTGAACGAGTTGCAGAATTGACGGCGCAGTTGAGGTGGT 1194
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558 CATTGGCAAAGTGGAAAAACGGTGAAACCGTGAACGAGTTGCAGAATTGACGGCGCAGTTGAGGTGGT 617

Qy   1195 AGTACCAAGAGGACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACA 1254
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   618 AGTACCAAGAGGACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACA 677

Qy   1255 TTTCTATGCCAGTCAGATGGCTCAACGGATCCGAGAACTCCGAGACATCCTGCCCAGGTTAAGCA 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   678 TTTCTATGCCAGTCAGATGGCTCAACGGATCCGAGAACTCCGAGACATCCTGCCCAGGTTAAGCA 737

Qy   1315 GCAGCATCAGAAGGACAGAGTGACAGTAACCAGGCCACGGACGAAGTGACCAGCCCCTCC 1374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   738 GCAGCATCAGAAGGACAGAGTGACAGTAACCAGGCCACGGACGAAGTGACCAGCCCCTCC 797

Qy   1375 CTGTCCCTTNGAGTCCAGGACAACACACGGGCAGAAATCGAGAGTGTGTCTCCCCGCAG 1434

```
Db    798  CTGTCCCTTCAGTCCAGGACAACACGGGCAGAAATCGAGAGTGTGCTCTCCCCGGCAG   857
           |||||||  ||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy   1435  GCCTGAGAATGAGTGGGAATCCGGGACACNTGGGCCCGGCGTGTAGATCAGGTTGCCCAC  1494
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    858  GCCTGAGAATGAGTGGGAATCCGGGACACCTGGGCCCGGCGTGTAGATCAGGTTGCCCAC   917

Qy   1495  TTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCACCCA  1554
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    918  TTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGATCTCTCAGCCCCAAACACCCACCCA   977

Qy   1555  ATTGGGCCCAACACTGTNTGCCCCTCGGGGTGTCAGAAATNTAGCCAAGGCACTTTTAA  1614
           |||||||||||||||| ||||||||||||||||||||||| ||||||||||||||||||
Db    978  ATTGGGCCCAACACTGTCTGCCCCTCGGGGTGTCAGAAATTCTAGCCAAGGCACTTTTAA  1037

Qy   1615  ACGTGGATTGTTTAAAGAAGTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTG  1674
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   1038  ACGTGGATTGTTTAAAGAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTG  1097

Qy   1675  GGAAGAAAAATAAAATTTCCTTCAGGTTTTAAAA  1708
           ||||||||||||||||||||||||||||||||||
Db   1098  GGAAGAAAAATAAAATTTCCTTCAGGTTTTAAAA  1131

RESULT 6
ABX73489
ID   ABX73489 standard; DNA; 1182 BP.
XX
AC   ABX73489;
XX
DT   18-MAR-2003  (first entry)
XX
DE   Human novel polynucleotide #317.
XX
KW   Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW   muscular disorder; respiratory disease; reproductive disorder;
KW   gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW   hyperproliferative disorder; inflammatory disease; allergic reaction;
KW   blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW   cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW   haemostatic; antiarteriosclerotic.
XX
OS   Homo sapiens.
XX
PN   US2002132753-A1.
XX
PD   19-SEP-2002.
XX
PF   17-JAN-2001; 2001US-00764864.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   28-JUN-2000; 2000US-0214886P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   22-AUG-2000; 2000US-0226868P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229444P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   08-SEP-2000; 2000US-0231413P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   27-SEP-2000; 2000US-0235834P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241809P.
PR   01-NOV-2000; 2000US-0246617P.
PR   17-NOV-2000; 2000US-0249299P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
XX
PA   (ROSE/) ROSEN C A.
PA   (RUBE/) RUBEN S M.
PA   (BARA/) BARASH S C.
XX
PI   Rosen CA, Ruben SM, Barash SC;
XX
DR   WPI; 2003-147444/14.
DR   P-PSDB; ABU55229.
XX
PT   New polypeptides and nucleic acids, useful in gene therapy for treating,
PT   inhibiting or preventing e.g. neural, immune system, muscular,
PT   respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT   renal disorders.
XX
PS   Claim 1; SEQ ID NO 327; 402pp; English.
XX
CC   The invention relates to human novel polypeptides and their associated
CC   polynucleotides. The polypeptides and polynucleotides are useful in gene
CC   therapy for treating, inhibiting or preventing neural disorders, immune
CC   system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC   and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC   nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC   gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC   (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC   heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC   renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC   leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC   appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC   related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC   infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC   human novel polynucleotides of the invention
XX
SQ   Sequence 1182 BP; 343 A; 310 C; 314 G; 215 T; 0 U; 0 Other;

Query Match             63.5%;  Score 1085.4;  DB 7;  Length 1182;
Best Local Similarity   98.9%;
Matches 1102;  Conservative    0;  Mismatches    6;  Indels    6;  Gaps    1;

Qy    601  CCCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGA   660
           | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     18  CGCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGA    77

Qy    661  AATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCCGCCATGAGC------TC   714
           ||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     78  AATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCGCGCCATGAGCCTGCAGTC   137
```

```
XX
AC   AAS26566;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human cDNA encoding a novel secreted protein, Seq ID 745.
XX
KW   Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW   cardiant; vasocropic; cerebroprotective; nootropic; neuroprotective;
KW   antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW   secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW   cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW   cerebral ischaemia; angiogenesis; nervous system disorder;
KW   Alzheimer's disease; infection; ocular disorder; corneal infection;
KW   wound healing; epithelial cell proliferation; skin ageing; food additive;
KW   preservative; antiproliferative.
XX
OS   Homo sapiens.
XX
PN   WO200155322-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001341.
XX
PR   31-JAN-2001;  2000US-0179065P.
PR   04-FEB-2000;  2000US-0180628P.
PR   24-FEB-2000;  2000US-0184664P.
PR   02-MAR-2000;  2000US-0186150P.
PR   16-MAR-2000;  2000US-0189874P.
PR   17-MAR-2000;  2000US-0190076P.
PR   18-APR-2000;  2000US-0198123P.
PR   19-MAY-2000;  2000US-0205515P.
PR   07-JUN-2000;  2000US-0209467P.
PR   28-JUN-2000;  2000US-0214886P.
PR   30-JUN-2000;  2000US-0215135P.
PR   07-JUL-2000;  2000US-0216647P.
PR   07-JUL-2000;  2000US-0216880P.
PR   11-JUL-2000;  2000US-0217487P.
PR   11-JUL-2000;  2000US-0217496P.
PR   14-JUL-2000;  2000US-0218290P.
PR   26-JUL-2000;  2000US-0220963P.
PR   26-JUL-2000;  2000US-0220964P.
PR   14-AUG-2000;  2000US-0224518P.
PR   14-AUG-2000;  2000US-0224519P.
PR   14-AUG-2000;  2000US-0225213P.
PR   14-AUG-2000;  2000US-0225214P.
PR   14-AUG-2000;  2000US-0225266P.
PR   14-AUG-2000;  2000US-0225267P.
PR   14-AUG-2000;  2000US-0225268P.
PR   14-AUG-2000;  2000US-0225270P.
PR   14-AUG-2000;  2000US-0225447P.
PR   14-AUG-2000;  2000US-0225757P.
PR   14-AUG-2000;  2000US-0225758P.
PR   14-AUG-2000;  2000US-0225759P.
PR   18-AUG-2000;  2000US-0226279P.
PR   22-AUG-2000;  2000US-0226681P.
PR   22-AUG-2000;  2000US-0226868P.
PR   22-AUG-2000;  2000US-0227182P.
PR   23-AUG-2000;  2000US-0227009P.
PR   30-AUG-2000;  2000US-0228924P.
PR   01-SEP-2000;  2000US-0229287P.
PR   01-SEP-2000;  2000US-0229343P.
PR   01-SEP-2000;  2000US-0229344P.
PR   01-SEP-2000;  2000US-0229345P.
PR   05-SEP-2000;  2000US-0229509P.
PR   05-SEP-2000;  2000US-0229513P.
PR   06-SEP-2000;  2000US-0230437P.
PR   06-SEP-2000;  2000US-0230438P.
PR   08-SEP-2000;  2000US-0231242P.
PR   08-SEP-2000;  2000US-0231243P.
PR   08-SEP-2000;  2000US-0231244P.
PR   08-SEP-2000;  2000US-0231413P.
```

```
Qy   715  TCACCTGATCCCTGGCCTGCTGTGTAGGTCTTTTCCCAGCTTCATCCAGCGC  774
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   138  TCACCTGATCCCTGGCCTGCTGTGTAGGTCTTTTCCCAGCTTCATCCAGCGC  197

Qy   775  AGTCCGACGCCGCTCCCAGCAGCGGTTACTGGGGCTGCTGCTCCTTTATGCAGGC  834
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   198  AGTCCCGACGCCGCTCCCAGCAGCGGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGC  257

Qy   835  TCCCGACCAGGAGATGTGTCAGGTGTTTATCCCGCCAGTGGCGCCATCATCG  894
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   258  TCCCGACCAGGAGATGTGTCAGGTGTTTATCCCGCCAGTGGCGCCATCATCG  317

Qy   895  CAAGAGGGGCAGCACATCAAACAGTCTCCGGTTGCCAGCGCGCCTCCATCAAGATTGC  954
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   318  CAAGAGGGGCAGCACATCAAACAGTCTCCCAAAGTTCGTATGGTTATCATCACTGACCGCCCAGAGGC  377

Qy   955  ACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCGCCCAGAGGC  1014
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   378  ACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACCGCCCAGAGGC  437

Qy   1015 CCAATTCAAGGCTCAGGGAAGAATCTATGCCAACCATACGTGTGCCAGCATCAGCAGCTCTTTTGGTCC  1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   438  CCAATTCAAGGCTCAGGGAAGAATCTATGCCAACCATACGTGTGCCAGCATCAGCAGCTCTTTTGGTCC  497

Qy   1075 CAAGGAGGAAGTGAAGCTGAAGACCCCATCAGCAGCAGTGCAGCTGAGGTGGGT  1134
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   498  CAAGGAGGAAGTGAAGCTGAAGACCCCATCAGCAGCAGTGCAGCTGAGGTGGGT  557

Qy   1135 CATTGGCAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACGGCGCAGTGAGTGGT  1194
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   558  CATTGGCAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACGGCGCAGTGAGTGGT  617

Qy   1195 AGTACCAAGAGACCAGACCCTGATGAGAACAACCAGTCATCGTGAAAATCATCGGACA  1254
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   618  AGTACCAAGAGACCAGACCCCTGATGAGAACAACCAGTCATCGTGAAAATCATCGGACA  677

Qy   1255 TTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCCTGCCCCAGGTTAAGCA  1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   678  TTTCTATGCCAGTCAGATGGCTCAACGGAACATCCTGCCCCAGGTTAAGCA  737

Qy   1315 GCAGCATCAGAGAGGGACAGAGTAACCAACAACGGGCACGGAGGAAGTGACCAGCCCTCC  1374
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   738  GCAGCATCAGAGAGGGACAGAGTAACCAACAACGGGCACGGAGGAAGTGACCAGCCCCTCC  797

Qy   1375 CTGTCCCTTNGAGTCCAGGACAACAACGGGCAGAAATCGAGAGTGTGCTCTCCCCGGCAG  1434
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   798  CTGTCCCTTCGAGTCCAGGACAACAACGGGCAGAAATCGAGAGTGTGCTCTCCCCGGCAG  857

Qy   1435 GCCTGAGAATGAGTGGGAATCCGGGACACNTGGGCCGGGCTGTGAGATCAGGTTTGCCCAC  1494
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   858  GCCTGAGAATGAGTGGGAATCCGGGACACCTGGGCCGGGCTGTGAGATCAGGTTTGCCCAC  917

Qy   1495 TTGATTGAGAAGATGTTCCAGTGAGGAACCCTGATCNTTCAGCCCAAAACACCCACCCA  1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   918  TTGATTGAGAAGATGTTCCAGTGAGGAACCCTGATCTCTCAGCCCCAAAACACCCACCCA  977

Qy   1555 ATTGGCCCAACACTGTNTGCCCTCGGGTGTCAGAAATTNTAGGCCAAGGCACTTTTAA  1614
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   978  ATTGGCCCAACACTGTTTGCCCTCGGGTGTCAGAAATTCTAGCCGCAAGGCACTTTTTAA  1037

Qy   1615 ACGTGGATTGTTTAAAGAAGAGTCTCCAGCCCCACCAAGAGGGTGGATCACACCTCAGTG  1674
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1038 ACGTGGATTGTTTAAAGAAGAGTCTCCAGCCCCACCAAGAGGGTGGATCACACCTCAGTG  1097

Qy   1675 GGAAGAAAATAAAAATTTCCTTCAGGTTTTAAAA  1708
              |||||||||||||||||||||||||||||||||
Db   1098 GGAAGAAAATAAAAATTTCCTTCAGGTTTTAAAA  1131
```

```
RESULT 7
AAS26566
ID   AAS26566 standard; cDNA; 1129 BP.
```

PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232357P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Barash SC, Ruben SM;

XX

DR WPI; 2001-488783/53.
DR P-PSDB; AAU16579.

XX

PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.

XX

PS Claim 1; SEQ ID NO 745; 980pp; English.

XX

CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match   62.6%; Score 1069.8; DB 4; Length 1129;
Best Local Similarity   98.7%; Pred. No. 1e-292;
Matches 1084; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

Qy 617 ACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCAGCAGGAAATAATGAAGAAAGTT 676
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCAGCAGGAAATAATGAAGAAAGTT 60

Qy 677 CGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGC------TCTCACCTGATCCCTGGC 730
            ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   61 CGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGCCTGCAGTCTCACCTGATCCCTGGC 120

Qy 731 CTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCGCCTCCC 790
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCGCCTCCC 180

Qy 791 AGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAGATG 850
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 AGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAGATG 240

```
Qy   851 GTGCAGTGTGTTTATCCCGCGCCAGGCAGTGGGCGCCATCATCGGCAAGAGAGGGCAGCAC  910
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTGCAGTGTGTTTATCCCGCGCCAGGCAGTGGGCGCCATCATCGGCAAGAGAGGGCAGCAC  300

Qy   911 ATCAAACAGCTCTCCCGGTTTGCCAGCGCCTTCATCAAGATTGCACCACCCGAAACACCT  970
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 ATCAAACAGCTCTCCCGGTTTGCCAGCGCCTTCATCAAGATTGCACCACCCGAAACACCT  360

Qy   971 GACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAG 1030
         |||||||||||||||||||:|:|||::|||||||||||||:||||||||||||||||||
Db   361 GACTCCAAAGTTCGTATGGKTATCATYACTGGACCGCCARAGGCCCAATTCAAGGCTCAG  420

Qy  1031 GGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTGGTCCCAAGGAGGAAGTGAAG 1090
         |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   421 GGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAG  480

Qy  1091 CTGGAGACCCACATACGTGCCAGCATCAGCAGCTGGCCCGGTCATTGGCAAAGGTGA 1150
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGGAGACCCACATACGTGCCAGCATCAGCAGCTGGCCCGGTCATTGGCAAAGGTGGA  540

Qy  1151 AAAACGGTGAACGAGTTGCAGAATTTGACGGCCAGCTGAGGTGGTAGTACCAAGAGACCAG 1210
         ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   541 AAAACGGTGAACGAGTTGCAGAATTTCACGGCCAGCTGAGGTGGTAGTACCAAGAGACCAG  600

Qy  1211 ACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATGCGACATTTCTATGCCAGTCAG 1270
         ||||| ||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   601 ACCCCYTGATGAGAACGACCAGGTCATCGTGAAAATCATGCCGACATTTCTATGCCCAGTCAG  660

Qy  1271 ATGGCTCAACGGAAGATCCGAGACATCCTGCCCCAGGTTAAGCACGCATCAGAGAGGGA 1330
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ATGGCTCAACGGAAGATCCGAGACATCCTGCCCCAGGTTAAGCACGCAGCAGAGAGGGGA  720

Qy  1331 CAGAGTAACCAGGCCCAGGCACGGGAGGAAGTGACCAGCCCCTCCCTGTCCCTTNGAGTCC 1390
         ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   721 CAGAGTAACCAGGCCCAGGCACGGGAGGAAGTGACCAGCCCCTCCCTGTCCCTTCGAGTCC  780

Qy  1391 AGGACAACAACGGGCAGAAATCAGAGAGTGTGCTTCTCCCCGGCAGGCCTGAGAATGAGTGG 1450
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGGACAACAACGGGCAGAAATCAGAGAGTGTGCTTCTCCCCGGCAGGCCTGAGAATGAGTGG  840

Qy  1451 NTGCCCCTCGGGGTGTCAGAAATTNTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAA 1510
         |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   841 CTGCCCCTCGGGGTGTCAGAAATTCTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAA  900

Qy  1511 TTCCAGTGAGGAACCCTGATCTNTCAGCCCAAAACACACCCAATTGGCCCAACACTGT 1570
         ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   901 TTCCAGTGAGGAACCCTGATCTCTCAGCCCAAAACACACCCAATTGGCCCAACACTGT  960

Qy  1571 NTGCCCCTCGGGGTGTCAGAAATTNTAGCGCAAGGAGGTGGATCACACCTCAGTGGGAAGAAATAAAAT 1630
         |||||||||||||||||||||||| ||||||||
Db   961 CTGCCCCTCGGGGTGTCAGAAATTCTAGCGCAAGGAGGTGGATCACACCTCAGTGGGAAGAAATAAAAT 1020

Qy  1631 GAAGCTCTTCCAGGCCCCACCCAAGAGAGGGTGGATCACACCTCAGTGGGAAGAAAATAAAAT 1690
         ||||||||||||||||||||||
Db  1021 GAAGCTCTTCCAGGCCCCACCCAAGAGAGGGTGGATCACACCTCAGTGGGAAGAAAAATAAAAT 1080

Qy  1691 TTCCTTCAGGTTTTAAAA 1708
         ||||||||||||||||||
Db  1081 TTCCTTCAGGTTTTAAAA 1098
```

RESULT 8
ABX73907

```
DR   P-PSDB; ABU55647.
XX
PT   New polypeptides and nucleic acids, useful in gene therapy for treating,
PT   inhibiting or preventing e.g. neural, immune system, muscular,
PT   respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT   renal disorders.
XX
PS   Claim 1; SEQ ID NO 745; 402pp; English.
XX
CC   The invention relates to human novel polypeptides and their associated
CC   polynucleotides. The polypeptides and polynucleotides are useful in gene
CC   therapy for treating, inhibiting or preventing neural disorders, immune
CC   system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC   and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC   nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC   gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC   (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC   heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC   renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC   leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC   appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC   related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC   infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC   human novel polynucleotides of the invention
SQ   Sequence 1129 BP; 318 A; 296 C; 302 G; 210 T; 0 U; 3 Other;

Query Match             62.6%;  Score 1069.8;  DB 7;  Length 1129;
Best Local Similarity   98.7%;  Pred. No. 1e-292;
Matches 1084;           Conservative    3;  Mismatches    5;  Indels    6;  Gaps    1;

Qy    617 ACTGTGAAGGGGCCATCGAGAATTGTTGCGAGGGCCGAGCAGGAAATAATGAAGAAAGTT 676
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACTGTGAAGGGGCCATCGAGAATTGTTGCGAGGGCCGAGCAGGAAATAATGAAGAAAGTT 60

Qy    677 CGGGAGGCCCATGAGAATGATGTGGCTGCCATGAGC-----TCTTCACCTGATCCCTGGC 730
          ||||||||||||||||||||||||||||||||||||     ||||||||||||||||||||
Db     61 CGGGAGGCCCATGAGAATGATGTGGCTGCCATGAGCGTCCACCTGATCCCTGATCCCTGGC 120

Qy    731 CTGAACCTGGCTGCTGTGTAGGTCTTTTCCCAGCTTCATCAGCGCAGGCTCCCGAGCAGGA 790
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGAACCTGGCTGCTGTGTAGGTCTTTTCCCAGCTTCATCAGCGCAGGCTCCCGCGCCTCCC 180

Qy    791 AGCAGGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATG 850
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGCAGGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATG 240

Qy    851 GTGCAGGTGTTTATCCCGCCCAGGCAGTGGGCCGCCATCATCGGCGCAAGAAGGGGCAGCAC 910
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGCAGGTGTTTATCCCGCCCAGGCAGTGGGCCGCCATCATCGGCGCAAGAAGGGGCAGCAC 300

Qy    911 ATCAAACAGCTCCTCCCGGTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCT 970
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATCAAACAGCTCCTCCCGGTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACCT 360

Qy    971 GACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGGCCAATTCAAGGCTCAG 1030
          |||||||||||||||::|||||||||||||||||||||||||||||||||||||||||||
Db    361 GACTCCAAAGTTCGTATGGXTATCATYACTGGACCGCCARAGGGCCAATTCAAGGCTCAG 420

Qy   1031 GGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAG 1090
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAG 480

Qy   1091 CTGGAGACCCACCATACGTGTGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGTGGA 1150
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGGAGACCCACCATACGTGTGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGTGGA 540

Qy   1151 AAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGTGGTAGTACCAAGGAGACCAG 1210
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AAAAGGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGTGGTAGTACCAAGAGACCAG 600

Qy   1211 ACCCCTGATGAGAACGACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAG 1270
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ACCCCTGATGAGAACGACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAG 660

Qy   1271 ATGGGCTCAACGGAGATCCGAGATCCTGGCCAGGTTAAGCAGCAGCATCAGAAGGGA 1330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ATGGGCTCAACGGAGATCCGAGATCCTGGCCAGGTTAAGCAGCAGCATCAGAAGGGA 720

Qy   1331 CAGAGTAACCAGGCCCAGGCACGGAGGAAGTGACCAGCCCTCCTTNGAGTCC 1390
          ||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    721 CAGAGTAACCAGGCCCAGGCACGGAGGAAGTGACCAGCCCTCCTTGTCCCTTCGAGTCC 780

Qy   1391 AGGACAACAACGGGACAGAAATCGAGAGTGTGCTCTCCCCGGAGGCCTGAGAATGAGTGG 1450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGGACAACAACGGGCAGAAATCGAGAGTGTGCTCTCCCCGGAGGCCTGAGAATGAGTGG 840

Qy   1451 GAATCCGGGACACNTGGGCGGCTGTAGATCAGTTGCCCACTTGATTGAGAAAGATG 1510
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    841 GAATCCGGGACACCTGGGCGGCTGTAGATCAGTTGCCCACTTGATTGAGAAAGATG 900

Qy   1511 TTCCAGTGAGGAACCCTGATCTNTCAGCCCCAAACACCCCAATTGGCCCAACACTGT 1570
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    901 TTCCAGTGAGGAACCCTGATCTCTCAGCCCCAAACACCCCAATTGGCCCAACACTGT 960

Qy   1571 NTGCCCCTCGGGGTGTCAGAAATNTAGCGCAAGGCACTTTAAACGTGGATTGTTTAAA 1630
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    961 CTGCCCCTCGGGGTGTCAGAAATTCTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAA 1020

Qy   1631 GAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTGGGAAGAAAATAAAAT 1690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTGGGAAGAAAATAAAAT 1080

Qy   1691 TTCCTTCAGGTTTTAAA 1708
          ||||||||||||||||
Db   1081 TTCCTTCAGGTTTTAAA 1098


RESULT 9
AAK91969
ID   AAK91969     standard; cDNA; 833 BP.
XX
AC   AAK91969;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human cDNA 5'-end sequence, SEQ ID NO: 429.
XX
KW   Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1130094-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-00114089.
XX
PR   08-JUL-1999;  99JP-00194486.
PR   11-JAN-2000;  2000JP-00118774.
PR   02-MAY-2000;  2000JP-00183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI   Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR   WPI; 2001-524255/58.
XX
PT   830 Primers useful for synthesizing full length cDNA clones and their use
PT   in genetic manipulation.
XX
PS   Claim 2; SEQ ID NO 429; 1380pp + Sequence Listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been isolated
```

RESULT 10

AAK93655
ID   AAK93655 standard; cDNA; 833 BP.
XX
AC   AAK93655;
XX
DT   06-NOV-2001 (first entry)
XX
DE   Human cDNA clone representative sequence, SEQ ID NO: 2115.
XX
KW   Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1130094-A2.
XX
PD   05-SEP-2001.
XX
PF   07-JUL-2000; 2000EP-00114089.
XX
PR   08-JUL-1999; 99JP-00194486.
PR   11-JAN-2000; 2000JP-00118774.
PR   02-MAY-2000; 2000JP-00183765.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI   Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR   WPI; 2001-524255/58.
XX
PT   830 Primers useful for synthesizing full length cDNA clones and their use
PT   in genetic manipulation.
XX
PS   Example 11; SEQ ID NO 2115; 1380pp + Sequence Listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. 830 cDNA molecules encoding a human protein have been isolated
CC   and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC   been determined. Primers for synthesising the full length cDNA are useful
CC   for clarifying the function of the protein encoded by the cDNA. The full
CC   length clones were obtained by construction of full length enriched cDNA
CC   libraries that were synthesised by the oligo-capping method. The primers
CC   enable the production of the full length cDNA easily without any special
CC   methods. The present sequence was used as the representative sequence
CC   from a human clone which was used in homology searches to identify the
CC   clone. Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in CD-ROM format directly from
CC   EPO
XX
SQ   Sequence 833 BP; 223 A; 225 C; 215 G; 166 T; 0 U; 4 Other;

Query Match   42.9%; Score 732.8; DB 4; Length 833;
Best Local Similarity   96.9%; Pred. No. 3.6e-197;
Matches 802; Conservative   0; Mismatches   14; Indels   12; Gaps   5;

[Sequence alignment data — two parallel blocks of Qy/Db nucleotide sequence rows]

Qy   299 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAA 358
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAA 60

Qy   359 AAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTTCCTCCGCTTGTAAGATGATCTTG 418
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTTCCTCCGCTTGTAAGATGATCTTG 120

Qy   419 GAGATTATGCATAAAGAGAGGCTAAGGACACCCAAAAACGGCTGACGAGGTTCCCTGAGATC 478
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAGATTATGCATAAAGAGAGGCTAAGGACACCCAAAAACGGCTGACGAGGTTCCCTGAGATC 180

Qy   479 CTGGCCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAG 538
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGGCCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAG 240

Qy   539 GTAGAGCAAGATACCGAGACAAAAAATCACCATCTCCTCGTTGCAAGAACCTTACCCTTTAC 598
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTAGAGCAAGATACCGAGACAAAAAATCACCATCTCCTCGTTGCAAGAACCTTACCCTTTAC 300

Qy   599 AACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCGAGGGCCGAGCAG 658
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCGAGGGCCGAGCAG 360

Qy   659 GAAATAATGAAGAAGTTCGGGAGGCCTATGAAGAATGATGTGGCTGCCATGAGCCTGCAG------ 712
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GAAATAATGAAGAAGTTCGGGAGGCCTATGAAGAATGATGTGGCTGCCATGAGCCTGCAG 420

Qy   713 TCTCACCTGATCCCTGGCCTGCGTGTCTTGGCTGCGTGTGAGGTCTTTTCCAGCTTCATCCAGC 772
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 TCTCACCTGATCCCTGGCCTGCGTGTCTTGGCTGCGTGTGAGGTCTTTTCCAGCTTCATCCAGC 480

Qy   773 GCAGTCCCGCCGCGCCTCCCAGCAGCGTTACTGGGCGTGCTCCCTATAGCTCCTTATGCAG 832
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GCAGTCCCGCCGCGCCTCCCAGCAGCGTTACTGGGCGTGCTCCCTATAGCTCCTTATGCAG 540

Qy   833 GCTCCCGAGCAGGAGATGTGCAGGTGTTTATCCCGCCCAGGCAGTGGGCCGCCCATCATC 892
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GCTCCCGAGCAGGAGATGTGCAGGTGTTTATCCCGCCCAGGCAGTGGGCCGCCCATCATC 600

Qy   893 GGCAAGGAGGGGCAGCACATCAAACAGCCTCCCGGTTGCAGCGCCTTCCATCAAGATT 952
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCAAGGAGGGGCAGCACATCAAACAGCTCTTCCAGCGCCTTCCAGCGCCTTCATCAAGATT 660

Qy   953 GCACCACCCGAAACACTTGACTTCGTATGGTTATCATCACTGGA-CCGCCAGA 1011
            |||||||||||||||||||||||||||||||||||||||||||
Db   661 GCACCACCCGAAACACCTGACTTCGTATGGTTATCATCACTGGACCGCCGCCAGA 720

Qy   1012 GGCCCAATTCAAGGCTCAGGGAGGAATCATGGCAAACTC-AAGGAGGAGAACTT--CTT 1068
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 NGCCCAATTCAAGGCTCANGGAGGAATCATGGCAAACTCAAGGAGGAGAACTTTCTTT 780

Qy   1069 TGGTCCCAAGGAGGAAGT--GAAGCTGGAGACCCACATACGTGTGCCA 1114
            ||||||||||||||||||||||||||||||||||||||
Db   781 GGTCCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAG 240

[Second alignment block on the page]

Qy   299 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAA 358
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAA 60

Qy   359 AAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTTGTAAGATGATCTTG 418
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTTGTAAGATGATCTTG 120

Qy   419 GAGATTATGCATAAAGAGAGGCTAAGGACACCCAAAAACGGCTGACGAGGTTCCCTGAGATC 478
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GAGATTATGCATAAAGAGAGGCTAAGGACACCCAAAAACGGCTGACGAGGTTCCCTGAGATC 180

Qy   479 CTGGCCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAG 538
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CTGGCCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAG 240

Query Match   42.9%; Score 732.8; DB 4; Length 833;
Best Local Similarity   96.9%; Pred. No. 3.6e-197;
Matches 802; Conservative   0; Mismatches   14; Indels   14; Gaps   5;

DR   P-PSDB; AAB11365.
XX
PT   Isolated polypeptide comprising an immunogenic portion of a lung tumor
PT   protein is used for detecting and monitoring progression of lung cancer
PT   in a patient.
XX
PS   Claim 1a; Page 258-259; 261pp; English.
XX
CC   This invention describes a novel isolated polypeptide (I) which
CC   comprising an immunogenic portion of a lung tumor protein or variant (P2)
CC   which have cytostatic activity. The polypeptides and polynucleotides are
CC   used in compositions and vaccines to inhibit the development of cancer,
CC   especially lung cancer, in a patient. Methods described in the invention
CC   can be used to monitor the progression of a cancer by carrying out the
CC   detection at subsequent time points and comparing the results from the
CC   different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC   are treated with P2, polynucleotides encoding P2 or antigen presenting
CC   cells expressing P2 and then administered to the patient to inhibit
CC   development of cancer
XX
SQ   Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;

Query Match   34.1%; Score 582.2; DB 3; Length 1740;
Best Local Similarity   66.2%; Pred. No. 2.8e-154;
Matches 858; Conservative   0; Mismatches 433; Indels   6; Gaps   1;

Qy   37 AGCCATCATGAAGCTGAATGCCACCAGTTGAGAACCATGCCCTGAAGGTCTCCTACAT 96
Db   402 AGCACTAGACAAACTGAATGGATTCAGTTAGAGAATTCACCTTGAAAGTAGCCTATAT 461

Qy   97 CCCCGATGAGCAGATAGCACAGGACCTGAGAATGGGCGCCGCAGCAGCCCCTTTCG 156
Db   462 CCCTGATGAAACGGGCCGCCCAGCAAACCCTTGCAGCAGCGGGGGCCCAGCAGCCAGT 521

Qy   157 GGGTCAGCCCCGCCAGGGCTCACCTGTGCAGCGGGGCCCAGCCAGCAGCAGCAAGT 216
Db   522 TGGGCAGAGGGGGTCCTCAAGGCAGGGGTCTCCAGGATCCGTATCCAAGCAGAAACCATG 581

Qy   217 GGACATCCCCTTCGGCGCTGCGGCTGCCACCCAGTATGTGGTTGCCATTATTGCCAAGGA 276
Db   582 TGATTTGCCTCTGCGACCTGCTGGTTCCCACCCATTTGTTGGTGAGCCATCATAGGAAAGA 641

Qy   277 GGGGGCCACCATCCGCAACATCCGGAACATCACCAACAGACAGCCCAGTCCAAGATAGAG 336
Db   642 AGGTGCCACCATTCGGAACATCACCAACATCACCAACAGACCAGTCTAAAATCGATGTCCACCGTAA 701

Qy   337 GGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGACTCACCCCTGAGGGCTGCTC 396
Db   702 AGAAAATGCGGGGCGGGCTGCTGAGAAGTCGATTACTATCCTCTCTACTCCTGAAGGCCACCTC 761

Qy   397 CTCCGCTTGTAAGATGATCTTTGAAGATGATCTTGGAGATTATGCATAAAGAGGGTAAGAGCACCAAAACGGC 456
Db   762 TGCCGGCTTGTAAGTCTATTCTGGAGATTATGCATAAGGAAGCTCAAGATAAAAATTCAC 821

Qy   457 TGACGGAGGTTCCCCTGAAGATCCTGGCCATAATAACTTTGTAGGGGGTCTCATTGGCAA 516
Db   822 AGAAGAGATCCCCTTGAAGATTTTAGCTCATAATAACTTTGTTGGACGTGCTTATTGGTAA 881

Qy   517 GGAAGGACGGAACCTGAACCTGAGAGAAGGTAGAGCAAGATACAGAAGATACCGAGACACAGAAGATCACCATCTCCTC 576
Db   882 AGAAGGAAGAAATCTTAAAAATTGAGCAAGACACAGAAGATCACGAGATATCTCC 941

Qy   577 GTTGCAAGACCTTACCCTTTACACCCTGAAACCCTTACAGGGGGCCATCGA 636
Db   942 ATTGCAGGAATTGACGCGTCTATAATCCAGAACGCCACTATTACAGTTAAGGCAATGTTGA 1001

Qy   637 GAATTGTTGCAGGGCCGAGCAGGACAGGAAATAATGAAGGCCTATGAGAATGA 696
Db 1002 GACATGTGCCAAAGCTGAGGCTGAGGAGGATCATGAGAGAGAAAATCAGGGAGTCTTATGAAAATGA 1061

Qy   697 TGTGGCTGCCATGA------GCTCTCACCTGATCCCTGGCTGCTGTAGG 750
Db 1062 TATTGCTTCTATGAAATCTTCAAGCACATTTAATTCCTGGATTAAATCTGAGAACGCCCTTGGG 1121

Qy   539 GTAGAGCAAGATACCGAGACAAAAATCACCATCTCCGTTGCAAGACCTTACCCTTTAC 598
Db   241 GTAGAGCAAGATACCGAGACAAAAATCACCATCTCCGTTGCAAGACCTTACCCTTTAC 300

Qy   599 AACCCTGAGAGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAG 658
Db   301 AACCCTGAGAGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAG 360

Qy   659 GAAATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGC------ 712
Db   361 GAAATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGCCTGCAG 420

Qy   713 TCTCACCTGATCCCTGGCTGCTGTGTAGGTCTTTCCCAGCTTCATCCAGC 772
Db   421 TCTCACCTGATCCCTGGCTGCTGTGTAGGTCTTTTCCCAGCTTCATCCAGC 480

Qy   773 GCAGTCCGCCGCCTCCCAGCAGCGTTACTGGGGCTTACTGGGGCTGCTCCCTATAGTCTCCTTTATGCAG 832
Db   481 GCAGTCCGCCGCCTCCCAGCAGCGTTACTGGGGCTTACTGGGGCTGCTCCCTATAGTCTCCTTTATGCAG 540

Qy   833 GCTCCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGGCCAGCAGTGGGGCCATCATC 892
Db   541 GCTCCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGGCCAGCAGTGGGGCCATCATC 600

Qy   893 GGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGTTGCCAGCGCCTTCATCAAGATT 952
Db   601 GGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGTTGCCAGCGCCTTCATCAAGATT 660

Qy   953 GCACCACCCGAAACACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGA-CCGCCAGA 1011
Db   661 GCACCACCCGAAACACACCTGACTTCAAAGTTCAAAGTTCATCATCACTGTTATCACTGGA 720

Qy 1012 GGCCCAATTCAAGGCTCAGGGAAGAATCTATTGGCAAACTC-AAGGAGGAGAACTT--CTT 1068
Db   721 NGCCAATTCAAGGCTCANGGAAGAATCTATGGCAAACTCAAAGGAGGAGAACTTTCTTT 780

Qy 1069 TGGTCCCAAGGAGGAAGT--GAAGCTGGAGACCCACATACGTGTGCCA 1114
Db   781 GGTCCCCAAGGAGGAAGTTGAAGCTTGGAGACCCACATACCTGTGCCA 828

RESULT 11
AAC66035
ID   AAC66035 standard; cDNA; 1740 BP.
XX
AC   AAC66035;
XX
DT   21-FEB-2001 (first entry)
XX
DE   Human lung cancer-associated cDNA antigen L523S.
XX
KW   Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW   vaccine; detection; ss.
XX
OS   Homo sapiens.
XX
PN   WO200061612-A2.
XX
PD   19-OCT-2000.
XX
PF   03-APR-2000; 2000WO-US008896.
XX
PR   02-APR-1999;   99US-00285479.
PR   17-DEC-1999;   99US-00466396.
PR   30-DEC-1999;   99US-00476496.
PR   10-JAN-2000;   2000US-00480884.
PR   22-FEB-2000;   2000US-0510576.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Fan L;
XX
DR   WPI; 2000-628399/60.

PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI Mcneill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
DR WPI; 2002-090513/12.
DR P-PSDB; ABB74997.
XX
PT Polynucleotides encoding lung tumor polypeptides, useful for treating
PT lung cancer or stimulating an immune response.
XX
PS Example 2; Page 330; 374pp; English.
XX
CC The present invention describes human lung tumour proteins. Human lung
CC tumour proteins and polynucleotides have cytostatic and immunostimulant
CC activities, and can be used in vaccine production. Compositions
CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC proteins, T cell populations, or antigen presenting cells that express
CC the lung tumour proteins are useful for treating lung cancer or
CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC ABB75070 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;

Query Match   34.1%; Score 582.2; DB 6; Length 1740;
Best Local Similarity 66.2%; Pred. No. 2.8e-154;
Matches 858; Conservative 0; Mismatches 433; Indels 6; Gaps 1;

Qy   37 AGCCATCATGAAGCTGAATGCCACCAGTTGAGAACCATGCCCGAAGGTCTCCTACAT 96
          |||| |||||||||||||| ||| ||||| |||||||| ||| |||||||||||||
Db   402 AGCCATTAGACAAACTGAATGATTTCAGTTAGAGAATTCACCTTAGAAGTAGCCTATAT 461

Qy   97 CCCCGATGAGCAGATAGCACAGGACCTGAGAATGGCGCCGAGCGGCCTTGGCTCTCG 156
          |||| || |||| ||| | || || |||||||||| ||||||||||| || |||||
Db   462 CCCTGATGAAACGGCGCGCCAGCGAAAACCCTGCAGCGGGGCCCCAGCCAGCAGCAAGT 521

Qy   157 GGGTCAGCCCCCCAGGGCTCACCTGTGCCAGCGGGGCCCCAGCCAGCAGCAGCAAGT 216
          ||||||| || ||||| ||||| |||| ||||| |||||| |||||||||||||||
Db   522 TGGGCAGAGGGGGCTCTTCAAGGCCAGGGGTCTCCAGGATCCGTATCCAAGCAGAAACCATG 581

Qy   217 GGACATCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGTGGCCATCATAGGAAAAGA 276
          || ||||||||||||||||||| || |||||||| ||||||| |||||| ||||||||
Db   582 TGATTTGCTGTCGCCCTGCTGTTCCCACCCAATTGTTGGAGCCATCATAGGAAAAGA 641

Qy   277 GGGGGCCACCATCCGCAACATCCAAACAGACCCCAGTCCAAGATAGACGTGCATAGGAA 336
          |||||||||||||||||| |||||||||||||| ||| ||||||||||||||||||||||
Db   642 AGGTGCCACCATTCGGAACATCCAACACCAGACCCAGTCTAAAATCCAGTCCACCGTAA 701

Qy   337 GGAGAACGGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCACCCTGAGGGCTGCTC 396
          |||||| |||| ||||||||||||||| ||||| ||||| || || ||||| |||||||
Db   702 AGAAAATGCGGGGCGTCGTGCTGAGAGTCGATTACTACTACATCCTCCTGAGGCACCTC 761

Qy   397 CTCCGCTTGTAAGATGATCTTGGAGATTATGCGAGAAGTAGCCAAGACACCCAAAACGGC 456
          | || ||||||||||||| |||| ||||||| || || ||||||| | |||||| ||
Db   762 TGCCGGCTTGTAAAGTCTATTCTGAGATTATGCATAAGGAAGCTCAAGATATAAAATTCAC 821

Qy   457 TGACGAGGTCCCCTGAAGATCCGGCCCATATAATAACTCTGTAGGGGGTCTCATTGGCAA 516
          ||| |||||| ||||||| ||| |||||| || ||||| ||| |||| ||||||||
Db   822 AGAAGAGATCCCCTTGAAGATTTTAGCTCATAATAACTTTGTTGGACGTCTTATTGGTAA 881

Qy   517 GGAAGGACCGAACCTGAAGAAGGTAGAGCAAGAATACCGAGACAAAATCACCATCTCCTC 576
          || ||| |||||||||||| |||||| | || || ||||| |||| |||||||||||||
Db   882 AGAAGGAGGAGAAATCTTAAAAGAAGGTGAGCAAGACACAGACACTAAAATTCAGGATATCTCC 941

Qy   577 GTTGCAAGACCTTACCCTTTACAACCCTGAGAGGGCCATCGA 636
          |||||| |||||| | ||| |||| |||||||||||||||
Db   942 ATTGCAGGAATTGACGCTGTATAATCCAGAACGCCACTATTACAGTTAAAGGCAATGTTGA 1001

Qy   637 GAATTGTTGCAGGGCCGGAGCACGCAGGAAATAATGAAGAAGTTCGGGAGGCCTATGAGAATGA 696
          |||| ||||||| ||||||||||||||||| ||| ||||||| |||||| |||| ||||||||
Db   1002 GACATGTGCCAAAGCTGAGGAGGAGAAATCATGAGAGAAATCATGAGAGAAATGA 1061

Qy   697 TGTTGGCTGCCATGA------GCTCTCACCTGATCCCTGGCTGCTGCTGTAGG 750
          ||| |||| |||||| || |||||||| ||| ||| | || || ||

RESULT 12
ABL49254
ID ABL49254 standard; cDNA; 1740 BP.
XX
AC ABL49254;
XX
DT 01-MAY-2002 (first entry)
XX
DE Human lung tumour L523S cDNA sequence SEQ ID NO:347.
XX
KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW immune response; ss.
XX
OS Homo sapiens.
XX
PN WO200200174-A2.
XX
PD 03-JAN-2002.
XX
PF 28-JUN-2001; 2001WO-US021065.
XX
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685696.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
XX
PA (CORI-) CORIXA CORP.
XX

Qy   751 TCTTTCCCAGCTTCATCCAGCGAGTCCGCCGCCTCCCAGCAGCGTTACTGGGGCTGC 810
             |||||||||||||||||||| ||| || || |||||||||| |||||| |||||
Db   1122 TCTGTTCCCACCCACTTCAGGGATGCCACCTCCCACCTCAGGGCCCCTCAGCCATGAC 1181

Qy   811 TCCCTATAGCTCCTTTATGCAGGCTCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGC 870
             || || ||| | | ||||| || || |||| |||| ||||||| ||| |||||||
Db   1182 TCCTCCCTACCCGCAGTTGAGCAATCAGAAACGGAGACTGTTCATCTGTTTATCCCAGC 1241

Qy   871 CCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGCAGCACATCAAACAGCTCTCCGGTT 930
           ||||| ||||||||||||||||||||||| ||||||||||||||||||| |||| ||
Db   1242 TCTATCAGTCGGTGCCATCATCGGCAAGCAAGCCAGCACATCAAGCAGCTTTCTCGCTT 1301

Qy   931 TGCCAGGCGCTCCATCAAGATTGCACCACCGCAAACACCTGACTCCAAAGTTCGTATGGT 990
             |||| ||| ||| ||||||| |||| | ||||||||| ||| || |||| |||||||
Db   1302 TGCTGGAGCTTCAATTAAGATTGCTCTCAGCCGGCACCAGAGTGCTAAAGTGAGGATGGT 1361

Qy   991 TATCATCACTGACCGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT 1050
           || ||||||| |||| | || ||||||||||| ||| || |||||||| |||||||
Db   1362 GATTATCACTGGACCACCAGAGCTCAGTTCAAGGCTCAGGGAAGAATTTATGGAAAAAT 1421

Qy   1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGAGAACCCCACATACGTGT 1110
          |||||| |||||| || || ||||||||||| |||| || |||| || || || ||||||
Db   1422 TAAAGAAGAAAAACTTTTGTTAGTCCTAAAGAACGAAGGTGAAACTTGAAGCTCATATCAGAGT 1481

Qy   1111 GCCAGCATCAGCAGCTGGCGCGGTCATTGGCAAAGGGTGGAAAAACGGTGAACGAGTTGCA 1170
             ||||||||||||| ||| ||||| ||| ||||||| ||||| || || ||||||||
Db   1482 GCCATCCTTTGCTGCTGGCCAGAGTTATTGGCAGAGGAGGCAAAACGGTGAATGAACTTCA 1541

Qy   1171 GAATTTGACGGCCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGAACGACCA 1230
          |||| || || |||||||| | |||| || ||||| |||| || |||||||| || |
Db   1542 GAATTTGTCAAGTGCAGAAATCATCGGACATTTCTATGCGACATTTCTATGCCAACTGAGATCATG 1601

Qy   1231 GGTCATCGTGAAAATCATCGGACATTCTATGCCGAGATCGTCCAACGGAAGATCCG 1290
          || || ||||||| |||||| |||| | |||| ||| |||| |||||||| | ||
Db   1602 AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTTGCCAGGTTGCCCAGAGAAAAATTCA 1661

Qy   1291 AGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAG 1327
          ||||||||| | |||| ||| || || | ||||||
Db   1662 GGAAATTCGACTCAGGTAAAGCAGCACCAACAACAG 1698

PI   Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR   WPI; 2002-583465/62.
DR   P-PSDB; ABP61917.
XX
PT   Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT   the polynucleotides, useful in pharmaceutical compositions such as
PT   vaccines and as markers to indicate the presence of lung cancer.
XX
PS   Claim 1; Page 337; 381pp; English.
XX
CC   The present invention describes isolated human lung carcinoma
CC   polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC   activity, and can be used in gene therapy and in vaccines. Compositions
CC   comprising (I) or (II) can be used for stimulating an immune response in
CC   a patient and for treating lung cancer in a patient. Oligonucleotides of
CC   (I) can be used for detecting the presence of a cancer in a patient, by
CC   obtaining a biological sample from the patient, contacting the biological
CC   sample with the oligonucleotide, detecting in the sample, an amount of
CC   polynucleotide that hybridies to the oligonucleotide and comparing the
CC   amount of polynucleotide that hybridises to the oligonucleotide to a
CC   predetermined cut-off value, and determining the presence of a cancer in
CC   the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC   vaccines. (I) is useful as a marker to indicate the presence or absence
CC   of a cancer such as lung cancer. ABQ92145 to ABQ2486 and ABP61866 to
CC   ABP61992 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;

Query Match   34.1%; Score 582.2; DB 6; Length 1740;
Best Local Similarity   66.2%;   Pred. No. 2.8e-154;
Matches 858; Conservative   0; Mismatches 433; Indels   6; Gaps   1;

Qy   37 AGCCATCATGAAGCTGAATGGCCACCAGTTGGAACCATGCCCCTGAGGTCTCCTACAT 96
              || |||| | | ||| ||||| ||||||||| || |||||||| ||||||| | |||
Db   402 AGCACTAGACAAACTGAATGGATTTCAGTTAGAGAATTCACCTTGAAAGTAGCCTATAT 461

Qy   97 CCCCGATGAGCAGATAGCACAGAGGACCTGAGAATGGGCGCGAGGGGGCTTTGCTCTCG 156
              || || | | ||| || ||| ||| ||| || ||| ||| |||| ||| ||| |
Db   462 CCCTGATGAAACGGCCGCCGCAGCAAAACCCTGCAGCGGGGCCCCGAGGTCGCCGGGGCT 521

Qy   157 GGGTCAGCCCCGCCAGGGCTCACCTGTGCCAGCGGGGCCAAGCAGCAGCAAGT 216
              ||| ||| || || ||||| | ||| | | | ||| | |||| ||||| |||
Db   522 TGGGCAGAGGGCGTCTCCAAGCAGGAGGGGTTCCAGGATCGTATCCAAGCAGAAAACCATG 581

Qy   217 GGACATCCCCTTCGGTCGGTGCCCACCCAGTGGAGATTATTGGGTGCCATTATTGGCAAGGA 276
              || | || || | ||| || || | ||||| | |||| ||||| | ||||||| ||||
Db   582 TGATTTGCCTCTGCGCCTGCTGCGTTCCCACCCAATTTGTTGGAGCCATCATAGGAAAAGA 641

Qy   277 GGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAAGATCAAGATGAGTGCATAGGAA 336
              || | || ||||| ||| ||| ||| | ||| ||||||||| ||||| || ||| ||||
Db   642 AGGTGCCCACCATTGGAACATCACCAACCCCAGTCTAAAATCGATGTCCACGGTAA 701

Qy   337 GGAGAACGGACGGTGCAGTGAAAAAGCCATCAGTGTGCACTCACCCCTGAGGGCTGCTC 396
              || || ||| ||||| ||||| ||||| ||||| || | |||||||| || ||| | ||
Db   702 AGAAAATGCGGGGCGTGCTGGAGAGTCGATTACTACATCCTCTCCTGAAGGCACCTC 761

Qy   397 CTCCGGCTTGTGTAAGATGATCTTGGAGATTATGTGCCATATAACTTTGTAGGGGTTCATTGGCAA 456
              || || ||| ||||||| ||||| || ||| ||| | | |||| |||| ||| | |||||||| ||
Db   762 TGCCGCTTGTAAGTCTATTCTGGAGAT+ATGCATAAGAAGGAAGCTCAAGATATAAATTCAC 821

Qy   457 TGACGAGTTCCCCTGAGAGTCCTGGCCCATGAGCATCTGGCCCATGAGGGCGTCTCATTGGCAA 516
              ||||||| |||| | ||| | |||| ||||| | ||||| || | ||| ||| |||||| ||
Db   822 AGAAGAGATCCCCTGAGATTTAGCTCATATAACTTTGTTGGACGTCTTTATTGGTAA 881

Qy   517 GGAAGGACGGAACCTGAGAAGGTAGAGTAGACGAGACAAAAATCACCATTCCTC 576
              || ||| | || | ||||||| | || || || || |||| |||| | |||
Db   882 AGAAGGAGAAATCTTAAAAAATTGAGCAAGACACAGACACTAAAATCACGATATCC 941

Qy   577 GTTGCAAGACCTTACCCTTTACACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGA 636
              ||||||| | || || |||| || | | |||| ||| || ||| ||| ||||| ||
Db   942 ATTGCAGGAATTGACGCGTGTATAATCCACCACTATTACAGTTAAGGCAATGTTGA 1001

Db   1062 TATTGCTTCTATGAAATCTTCAAGCACATTTAATTCCTGGATTAAATCTGAACGCCCTTGGG 1121
              | |||| |||| ||||| | ||||| |||| | |||||||| |||||||||| || ||||||
Qy   751 TCTTTTCCAGCTTCATCCAGCAGCAGTCCGCGCTCCCAGCAGCGGTTACTGGGGCTGC 810

Db   1122 TCTGTTCCCACCCACTTCAGGGATGCCACCTCCACCTTCAGGGCCCCCTTCAGCCATGAC 1181
              ||| || |||| || |||||| ||| | ||||||||| || ||| || ||||| |
Qy   811 TCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGTGCAGGTGTTTATCCCCGC 870

Db   1182 TCCTCCCTACCCGCAGTTTGAGCAATCAGAAACGGAGACTGTTCATCTGTTTATCCCAGC 1241
              ||| ||| | ||||||||||||||| ||||||| |||| ||| || | ||| ||||||| |
Qy   871 CCAGGCAGTGGGGCCATCATCGGCAAGAAGGGGCAGCACATCAAACAGCTTCTCCCCGTT 930

Db   1242 TCTATCAGTCGGTGCCATCATCGGCAGCGACGAGGGGCACGGGGCCAGCACATCAAGCAGCTTTCTGCTT 1301
              |||||||| ||||||||||||||| || || |||||| || | |||| ||| |||||| |||| ||||
Qy   931 TGCCAGCGCCTCCATCAAGATTGCACACCTGACTCCAAAGTTCGTATGGT 990

Db   1302 TGCTGGAGTTCAATTAAGATTGCTCTCCAGCGGAGCCACCAGATGCTAAAGTGAGGATGGT 1361
              ||||||| ||||| |||||||||| ||||| || || || || || |||||||||||||||
Qy   991 TATCATCACTGGACGGCCAGAGCGCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT 1050

Db   1362 GATTATCACTGGACCACCAGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGGAAAAAT 1421
              ||||||||||||| ||||||||| ||||| |||||||||||||||||| || || |||||
Qy   1051 CAAGGAGAACTTCTTTGGTCCCAGGAGGAAGTGAAGCTGGAGACCCACATACGTGT 1110

Db   1422 TAAAGAAGAAACTTTGTTAGTCTAAAGAAGAGGTGAAACTTGAAGCTCATATCAGAGT 1481
              |||||||||| ||||||| |||| |||||||| |||| ||||||||| || ||||||
Qy   1111 GCCAGCATCAGCAGCTGCGCCGGGTCATTGCCAAAAGTGGAAAAACGGTGAACGAGTTGCA 1170

Db   1482 GCCATCCTTTGCTGCTGCGAGAGTTATTGGAAAAGAGGCAAAACGGTGAATGAACTTCA 1541
              |||| |||||||||||| || |||||||||||| ||| ||| ||||| |||| ||||
Qy   1171 GAATTTGACGCGGCAGCTGAGGTGGTGCAGGTACCAGAGAACGACCA 1230

Db   1542 GAATTTGTCAAGTGCAGAAGTTGTTGTCCCCTCGTGACCAGACACCTGACGAATGACCA 1601
              ||||||| ||||||||||||||||||| | || | ||||| ||||| |||||||||||
Qy   1231 GGTCATCGTGAAAATCATCGGACATTCTATGCCAGTCAGATGGCTCAACGGAAGATCCG 1290

Db   1602 AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTTGCCAGTTGCCCAGAGAGAAAATTCA 1661
              ||||||||||||||||| ||| || ||||||||||||||||| || ||||||||||||||
Qy   1291 AGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAG 1327

Db   1662 GGAAATTCGACTCAGGTAAAGCAGCTAAAGCGCAGCACCAACAACAG 1698

RESULT 13
ABQ92440
ID   ABQ92440 standard; cDNA; 1740 BP.
XX
AC   ABQ92440;
XX
DT   07-OCT-2002 (first entry)
XX
DE   Human lung cancer associated cDNA sequence SEQ ID NO:347.
XX
KW   Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine; gene;
KW   ss.
XX
CS   Homo sapiens.
XX
PN   WO020247534-A2.
XX
PD   20-JUN-2002.
XX
PF   30-NOV-2001; 2001WO-US047576.
XX
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
PR   28-JUN-2001; 2001US-00897778.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI   Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;

```
PR   21-AUG-2000; 2000US-00643597.
PR   15-SEP-2000; 2000US-00662786.
PR   09-OCT-2000; 2000US-00685696.
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI   Mcneill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
PI   Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX
DR   WPI; 2002-090513/12.
DR   P-PSDB; ABB75054.
XX
PT   Polynucleotides encoding lung tumor polypeptides, useful for treating
PT   lung cancer or stimulating an immune response.
XX
PS   Claim 1; Page 370; 374pp; English.
XX
CC   The present invention describes human lung tumour proteins. Human lung
CC   tumour proteins and polynucleotides have cytostatic and immunostimulant
CC   activities, and can be used in vaccine production. Compositions
CC   comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC   proteins, T cell populations, or antigen presenting cells that express
CC   the lung tumour proteins are useful for treating lung cancer or
CC   stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC   ABB75070 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 1743 BP; 527 A; 406 C; 418 G; 392 T; 0 U; 0 Other;

Query Match             34.1%;  Score 582.2;  DB 6;  Length 1743;
Best Local Similarity   66.2%;  Pred. No. 2.8e-154;
Matches   858;  Conservative   0;  Mismatches  433;  Indels   6;  Gaps   1;

Qy         37 AGCCATCATGAAGCTGAATGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACAT     96
              ||  ||||       ||||||||||||||||||||||||||||||   |||||   |||
Db        402 AGCACTAGACAAACTGAATGAATGATTTCAGTTAGAGAATTTCAGTTGAAAGTAGCCTATAT   461

Qy         97 CCCCGATGAGCAGATAGCACAGGACCTGAGAATGGCGGCCGCAGTGGGCCTTGGCTCTCG    156
              |      |||||||||||||||||||||  |||||||||||||||||||||||||  |||
Db        462 CCCTGATGAAACGGCCGCCCCAGCAAAACCCCTGCGAGCGGGGCCCCCAGCAGCAGCAAGT    521

Qy        157 GGGTCAGCCCCCGCCAGGGCTCACCTGTGGCAGCAGGGGTCTCCAGGATCCGTATCCAAGCAGAAACCATG    216
              |||   |||| |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        522 TGGGCAGAGGGGCTCCTCAAGGCAGGGGTTCCACCCAGTATGTGGGTGCCATTATTGGAGCGCATCATAGGGA    581

Qy        217 GGACATCCCCTTCGGCTCCTGCTGGTCCTGGTGCCACCCAGTAGATGTGGGTGCCATTATTGGAGCATCATAGGAAAAGA    276
              |||  ||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 TGATTTGCCTCTGCGGCCTGCTGGTTCCACCCCAATTTGTTGTTGGAGCGCATCATAGGAGAAAGA    641

Qy        277 GGGGGCCACCATCCGCAACATCCGAACATCACAAAACAGACCCAGTCAAGATAGACGTGCATAGGAA    336
              ||     |||||||  ||||| ||||| |||||||||||||  |||||| |||   ||| |||||
Db        642 AGGTGCACCATTCGGAACATCACCAAACAGACCCAGTCTAAAATCGATGTCCACCGTAA    701

Qy        337 GGAGAACGACAGTGCAGTCGAGCTGAAAAAGCCATCAGTGTGCACCCTCGAGGGCTGCTC    396
              |||||||||||||||||||||||||||||||||||||   ||| |||||||||||||||
Db        702 AGAAAATGCGGGGCGTGCTGCTGGAAGTGCAGGAAGTCGATTACTATCCTCTCACTACTCCTGAAGGCACCTC    761

Qy        397 CTTCCGCTTGTAAGATGATCATCTTGGAGATTATGCATAAGAGCGAGCAGCTAAGGATCCATAAAACGGGC    456
              |||     |||||   |||||||| |||||||||   |||||||| ||| ||||| |||  ||||||
Db        762 TGCGGGCTTGTAAGTCTATTCTTCTGGAGATTATGCATAAGAGGAAGCTCAAGATCTAAAATCACTCAC    821

Qy        457 TGACGAGGTTCCCTGAGATCCTGGCCCATAATAACTTTGTAGGAGGCGTCCATTCATTGGCAA    516
              |||||||||| |||||||||||||| ||||||||||  |||||   |||||||||||||  |||
Db        822 AGAAGAGATCCCCTTGAAGATTTTAGCTCATAATAACTTTGTTGGACGTCTTATTGGTAA    881

Qy        517 GGAAGGACGGAACCTGAAGAGTGAGAGCAGGTAGAGCAAGAAATCACCATCTCCTC    576
              |||| ||||   |||||||||||||||||||| |||||||||||||||  ||||||
Db        882 AGAAGGAGGAAGAAATCTTAAAAAAATTGAGCAAGACACAACAGACACTATAATCACGATATCCC    941

Qy        577 GTTGCAAGACCTTACCCTTACAACCCTGAAGGACCATCACTGTGAAGGGGCCATCGA    636
```

```
Qy        637 GAATTGTTGCAGGGGCCGAGCGAGCAGAAATAATGAAGAAAGTTCGGAGGCCTATGAGGATGA    696
              ||||||||||| |||||| ||||| |||||  |||  ||||||||||| ||||| ||||||||
Db       1002 GACAGTGTGCCAAAGTGCCAAAGTCATGAGGAGGATCATGAAGAAAATCAGGAGTCTTATGAAAATGA   1061

Qy        697 TGTGGGCTGCCATGA------GCTCTCACCTGATCCTGGCCTGGATTAAATCTGGCTGCTGTAGG    750
              | |||  ||||||||      |||||  ||||||||||||  || ||||||||||  ||||||||
Db       1062 TATTGCTTCTATGAATCTTCAAGCACACATTTATTCCTGGACACATCTGAAACGGCCTTGGG   1121

Qy        751 TCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCTCCCAGCAGCAGCGTTACTGGGGCTGC    810
              |||   ||||||  |||| |||||||  ||||||||||||| |||| |||||||||||||
Db       1122 TCTGTTCCCACCCACTTCAGGGGATGCCACCTCCACCTTCCAGGGCCCCCTTCAGCCATGAC   1181

Qy        811 TCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGTGCAGGTGTTTATCCCCGC    870
              ||||  ||| ||||| ||||| ||||  |||||  |||| ||||||| |||||||||||
Db       1182 TCCTCCCTACCCGCAGTTTGAGCAATCAGAAAACGGAGACTGTTCATCGTGTTTATCCCAGC   1241

Qy        871 CCAGGCAGTGGGCGCCATCATCGGCAAGAGAGGGCAGCACATCAAAACAGCTCTCCCGGTT    930
              |||||||||| |||| ||||||||| |||||||||| |||| ||||||||||||  ||| |
Db       1242 TCTATCAGTCGGTGCCATCATCGGCCAGCACATCAAGCAGCTTTCTCGCTT   1301

Qy        931 TGCCAGCGCCTCCATCAAGATTGCCACCACCGAAACACCTGACTCCCAAAGTTCGTATGGT    990
              ||||||  |||| |||||| |||| ||||||||||||||| |||   ||||||||||||||
Db       1302 TGCTGGAGCTTCAATTAAGATTGCTCCAGCCGAAGCCACCAGATGCTAAAGTGAGGATGGT   1361

Qy        991 TATCATCACTGACCGCCAGAGCGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT   1050
              ||| |||||||||||  |||| ||||| |||||||||| |||| ||  |||  |||||||
Db       1362 GATTATCACTGACCAGCCAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGGAAAAT   1421

Qy       1051 CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGT   1110
              ||||||| |||||||| ||||   ||||||||||||||||| |||| |||||||||||||
Db       1422 TAAAGAAGAAAACTTTGTTAGTCCTAAAGAACGAGAGTGAAAACTTGAAGCTCATATCAGAGT   1481

Qy       1111 GCCAGCATCAGCAGCTGCGCCGGTCATTGGCCAAAGTGGAAAAACGGTGAACGAGTTGCA   1170
              |||||||||||||||||| || |||||||  || ||||||||| ||||||||||||||||
Db       1482 GCCATCCTTTGCTGCTGGCAGAGTTATTGCAAGTGCTGGCAAAAGCGGTGAATGAACTTCA   1541

Qy       1171 GAATTTGACGGCAGCTGAGGTGTAGTACCAAGAGACCAGACCCTGATGAGAACCACCA   1230
              |||||| |||||| |||||| ||||  |||||| |||||||||||||||||||||||||
Db       1542 GAATTTGTCAAGTGCAGAAGTGTTGTCCCCTCGTGACCAGACACCTGATGAGAATGACCA   1601

Qy       1231 GGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGTCAACCGGAAGATCCG   1290
              |||| |||| || |||| |||| |||||||||||||| ||   | |||  |||||||
Db       1602 AGTGGTTGTCAAAAATAACTGGTCACTTCTATGCTTGCCCAGTTGCCCAGAGAAAAATTCA   1661

Qy       1291 AGACATCTGGCCCAGGTTAAGCAGCAGCATCAGAAG   1327
              |||||| |||||||||||||||||| |||||||||
Db       1662 GGAAAATTCTGACTCAGGTAAAGCAGCAGCACCAAACAACAG   1698
```

RESULT 14
ABL49299
ID   ABL49299 standard; cDNA; 1743 BP.
XX
AC   ABL49299;
XX
DT   01-MAY-2002 (first entry)
XX
DE   Human lung tumour L523S recombinant polynucleotide SEQ ID NO:450.
XX
KW   Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW   immune response; ss.
XX
OS   Homo sapiens.
XX
FN   WO200200174-A2.
XX
PD   03-JAN-2002.
XX
PF   28-JUN-2001; 2001WO-US021065.
XX
PR   28-JUN-2000; 2000US-00606421.
PR   02-AUG-2000; 2000US-00630940.

PR   28-JUN-2000; 2000US-00606421.
PR   02-AUG-2000; 2000US-00630940.
PR   21-AUG-2000; 2000US-00643597.
PR   15-SEP-2000; 2000US-00662786.
PR   09-OCT-2000; 2000US-00685696.
PR   12-DEC-2000; 2000US-00735705.
PR   07-MAY-2001; 2001US-00850716.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI   Mcneill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
PI   Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX
DR   WPI; 2002-090513/12.
DR   P-PSDB; ABB75053.
XX
PT   Polynucleotides encoding lung tumor polypeptides, useful for treating
PT   lung cancer or stimulating an immune response.
XX
PS   Claim 1; Page 367; 374pp; English.
XX
CC   The present invention describes human lung tumour proteins. Human lung
CC   tumour proteins and polynucleotides have cytostatic and immunostimulant
CC   activities, and can be used in vaccine production. Compositions
CC   comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC   proteins, T cell populations, or antigen presenting cells that express
CC   the lung tumour proteins are useful for treating lung cancer or
CC   stimulating an immune response. ABL49959 to ABL49300 and ABB74946 to
CC   ABB75070 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 1743 BP; 527 A; 406 C; 418 G; 392 T; 0 U; 0 Other;

Query Match   34.1%; Score 582.2; DB 6; Length 1743;
Best Local Similarity   66.2%;
Matches 858; Conservative   0; Mismatches 433; Indels   6; Gaps   1;

Qy   37 AGGCCATCATGAAGCTGAATGGCCACCAGTTGAGAACCATGCCGAGGTCCTCACAT 96
                    ||||||| |||||| |||| ||| ||||| |||||| |||||
Db   402 AGCACTAGACAAACTGAATGGATTTCAGTTGAAGTAGCCTATAT 461

Qy   97 CCCCGATGAGCAGATAGCACAGGGACCTGAGAATGGCGCCGAGAGTCCCGGGGGCT 156
                ||| |||||| || |||| || |||||| ||| || || || |||||||
Db   462 CCCTGATGAAACGGCCGCGCCCAGCAAACCCTTGCAGCGGGGCCCGAGGTCGCGGGGGCT 521

Qy   157 GGGTCAGCCCGCCAGGGCTCACCTGTGCAGCGGGGCCCAGCCAAGCAGCAGCAAGT 216
             ||| ||| | ||||| || ||| |||| |||| |||| || |||| |||| |
Db   522 TGGGCAGAGGGGGCTTCCTCAGGCGCAGGGGTCTCCAGGATCCGTATCCAAGCAGAGAAACCATG 581

Qy   217 GGACATCCCCCTTCGGCTCCTGCTGCTCGTGCCCACCCAGTATGTGGTGCCATTATTGCCAGGA 276
             || ||||| | ||||| |||| |||||||| |||| ||| |||| |||| || |||||
Db   582 TGATTTGCTTCGTCTGCGCCTGCGTGCTCGTGTTCCCAATTTGTTGGAGGCCATCATAGAGCTGCAGGAAAGA 641

Qy   277 GGGGGCCCACCATCCGCAACATCCTGGAACATCAACATCAGACCAGACCCAGTGTATAAATT 336
             |||| |||||||| | |||||||| |||||||||||||||| || ||||| | | ||||
Db   642 AGGTGCCCACCATTGGAACATCAACATCAGACCCAGTCTAAATGTCTCACCGTAA 701

Qy   337 GGAGAACGACAGTGGCAGCTGAAAAAAGCCATCAGTGTGCACTCCACCCTGAGGGCTGCTC 396
             | |||| | |||| ||| ||||||| || || | | | |||| || |||| ||||||
Db   702 AGAAAATGCGGGGCTGCTGAAGTCGATTCGATTACTATCCTCTTACTCCTGAAGGCCACCTC 761

Qy   397 CTCCGCTTGTAAGATGATCTTGGAGATGATCTTGGAGATTTGCATAAGAGGTCATCATAAGAGGACCAAAAACGGC 456
             | | ||||||||| | || |||||||| |||| ||||| | ||||| | | || | ||| |||||||||||| ||
Db   762 TGCGGCTTGTAAGTCTATTCTGGAGATTATGCATAATATGCATAAGTCAAGATCAAGATAAAATTCAC 821

Qy   457 TGACGAGGTTCCCCTGAAGATCCTGAGATCCCCATAATAACTTTGTAGGGCGTCTCATTGGCAA 516
              ||| ||||| ||||||||||| | |||| |||||||||||||| || ||||| |||| ||
Db   822 AGAAGAGATCCCCTTGAAGATTTTAGCTCATAATAACTTTGTTGGAGCGTCTTATTGGTAA 881

Qy   517 GGAAGGACGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAATCACCATCCTC 576
             ||||||| | || ||| ||||| | | |||| || ||||||||| ||| ||| ||
Db   882 AGAAGGAGAAATCTTAAAAAATTGAGCAAGACACACAGCCAAGACACACTAAAATCACGATATCTCC 941

Db   942 ATTGCAGGAATTGACGCTTGTATAATCCAGAACCACTATTACAGTTAAAGGCAAGTGTTGA 1001
                ||| ||| || ||| |||||||||| || || ||| || |||| |||
Qy   637 GAATTGTTGCAGGGCCAGCCGAGCAGGAAATATGAGAAAGTTCGGGAGGCCTATGAGAATGA 696

Db   1002 GACATGTGCCAAAGCTGAGCGAGGAGAGATCATGAGAGAAATCAGGGAGTCTTATGAAAATGA 1061
               ||||||| | ||||||||| ||| ||||| ||| ||||| | | | |||| |||||
Qy   697 TGTGGCTGGCCATGA------GCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCGTGTAGG 750

Db   1062 TATTGCTTCTATGAATCTTCAAGCACATTTAATTCCTGGATTAAAATCTGAACGCCTTGGG 1121
              | ||| || || | ||| ||| | ||| || |||| | || | |||| || ||||||
Qy   751 TCTTTTCCAGCTTCATCCAGCAGTCCCGCCTTCCCAGCAGCGTTACTGGGGCTGC 810

Db   1122 TCTGTTCCCACCCACTTCAGGGATGCCACCTCCCACCTCAGGCGCCCCTTCAGCCATGAC 1181
             ||| | | ||| ||| | ||| ||||||| |||| || ||| | ||||||||| ||
Qy   811 TCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGC 870

Db   1182 TCCTCCCTACCCGGCCAGTTGAGCAATCAGAAAACGGAGACTGTTCATCTGTTTATCCCAGC 1241
             ||| | || | ||| | || || | || |||| || ||||| ||| |||||||||| ||
Qy   871 CCAGGCAGTGGGCGCCATCATCGGCCAAGAAGGGCAGCACATCAAACAGCTCTCCCGGTT 930

Db   1242 TCTATCAGTCGGTGCCATCATCCGCAGCAGGGCCAGCACATCAAGCAGCTTTCTGCTT 1301
              | || ||| | |||||||||| |||| ||| |||||||||||| |||| ||| |||
Qy   931 TGGCCAGCGCCTCCATCAAGATGCACCACCGCAAAACCTGACTCCAAAGTTCGTATGGT 990

Db   1302 TGCTGGAGCTTCAATTAAGATTGCTCCAGCGAAGCACCAGATGCTAAAGTGAAGGATGGT 1361
             || | || ||| ||| ||||| || ||||| ||| |||| || | ||||| | |||||
Qy   991 TATCATCACTGACACGCCAGAGCCCAATTCAAGGCTCAGGGAAGAATCTATGCCAAACT 1050

Db   1362 GATTATCACTGACACCACCAGAGGCTCAGTTCAAGGCTCAGGGAAGAATTTATGAAAAAT 1421
              | | |||||||||| || ||| |||| | ||||||||||||||||| ||||| ||| |
Qy   1051 CAAGGAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTTGAGACGACCCACATACGTGT 1110

Db   1422 TAAAGAAGAAACTTTGTTAGTCCTAAAGAAGAGAGGTGAAACTTGAAGCTCATATCAGAGT 1481
              |||| || || |||| | |||| |||||| || || |||| |||||| ||||||||| |
Qy   1111 GCCAGCATCAGCAGCTGGCCGCGTCATTGGCCAAAGGTGGAAAAACGGTGAACGAGTTGCA 1170

Db   1482 GCCATCCTTTGCTGCTGGCAGAGTTATTGGAAAAGGAGGCAAAAACCGGTGAATGAACTTCA 1541
             ||||| || ||||| |||| | | ||||| ||| || || |||||| ||||||| || ||||
Qy   1171 GAATTTGACGGCCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGAACGACCA 1230

Db   1542 GAATTTGTCAAGTGCGAGAAGTTGTTGTTCCCTCGTGACAGACACCTGACAGCAATGACCA 1601
             ||||||| | || |||| || | | | || || ||||| |||| || ||||||||
Qy   1231 GGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGTCAGATGGCTCAACGGAGATCCG 1290

Db   1602 AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTTGCCAGTTGCCCAGAGAAAAATTCA 1661
             | || || || |||| | ||| || |||||||| | ||||| | | ||||||||||| |
Qy   1291 AGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAG 1327

Db   1662 GGAAATTCTGACTCAGGTAAAGCAGCAGCACCAACACAG 1698
             | || ||||| || |||| ||||||||||| || ||
Qy   ... AGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAG 1327

RESULT 15
ABL49297
ID   ABL49297 standard; cDNA; 1743 BP.
XX
AC   ABL49297;
XX
DT   01-MAY-2002 (first entry)
XX
DE   Human lung tumour L523S recombinant polynucleotide SEQ ID NO:447.
XX
KW   Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW   immune response; ss.
XX
OS   Homo sapiens.
XX
PN   WO200200174-A2.
XX
PD   03-JAN-2002.
XX
PF   28-JUN-2001; 2001WO-US021065.
XX

```
Qy    577  GTTGCAAGACCTTACCCTTACAACCCTGAGAGGACCATCACTGTGAAGGGGGGCCATCGA  636
                ||||| || ||   |||| ||||| |  ||| |||| ||  ||| |  ||||   ||
Db    942  ATTGCAGGAATTGACGCTGTATAATCCAGACGCACTATTACAGTTAAAGCAATGTTGA  1001

Qy    637  GAATTGTTGCAGGGCCCAGCCAGGCAGGAAATAATGAAGAAAGTTCGGGAGGCCTATGAGAATGA  696
                |||| |  ||| ||||||| | ||  || ||| ||  |||  |||| | |||| |||||||
Db   1002  GACATGTGCCAAAGCTGAGGAGGAGATCATGAAGAAAATCAGGGAGTCTTATGAAAAATGA  1061

Qy    697  TGTGGCTGCCATGA-----GCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAGG  750
                || | || ||||    |||  |||   |||| || ||   |||  | || |||| |
Db   1062  TATTGCTTCTATGAAATCTTCAAGCACATTAATTCCTGGATTAAATCTGAATCGGCCTTGGG  1121

Qy    751  TCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCGCCTCCCAGCAGCGTTACTGGGGCTGC  810
                |  ||||||||| |||||| ||   |||  | ||||  || ||  || ||| ||
Db   1122  TCTGTTCCCACCCACTTCAGGGGATGCCACCTCCCACCTCAGGGCCCCCTTCAGCCATGAC  1181

Qy    811  TCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGGATGGTGCAGGTGTTTATCCCCGC  870
                |||  || | ||| |||||| | ||  |  |||  |||  |  ||  ||| ||| ||
Db   1182  TCCTCCCTACCCGCAGTTTGACAATCAGAAAACGGAGACTGTTCATCTGTTTATCCCCAGC  1241

Qy    871  CCAGGCAGTGGGCGCCCATCATCGGCAAGAAGGGGCAGCACATCAAACAGCTCTCCCGGTT  930
                ||||||| ||  ||  |||||||||||||||| || ||  |||  ||||| ||  |||
Db   1242  TCTATCAGTCGGTGCCATCATCGGCAAGCAGGGCCAGCACACTCAAGCAGCTTTCTCGCTT  1301

Qy    931  TGCCAGCGCCTCCATCAAGATTGCACCACCCGAAAACACCTGACTCCAAAGTTCGTATGGT  990
                ||||||| || | || ||||| ||  | ||||||| ||  ||  ||  ||| | |||||
Db   1302  TGCTGGAGCTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTAAAGTGGAGGATGGT  1361

Qy    991  TATCATCACTGACCGCGCAGAGCGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAACT  1050
                |||||||||||| ||||| ||| |||| ||||||||| |||||| |||| |||| |||||
Db   1362  GATTATCACTGACCGACCACCAGAGGCTCAGGCTCAAGGCTCAGGGAAGAATTTATGGAAAAAT  1421

Qy   1051  CAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGT  1110
                |||||||||||| ||| ||||||  ||| || ||| || ||| ||||| ||| ||| |
Db   1422  TAAAGAAGAAAACTTTGTTAGTCCTAAAGAAGAGGTGAAACTTGAAGCTCATATCAGAGT  1481

Qy   1111  GCCAGCATCAGCAGCTGCGCCGGTCATTGGCCGGGTCAAAGGTGAAAAACGGTGAACGAGTTGCA  1170
                |||||||||| || ||  ||  ||| |||| |||| ||||| || |||  |||  ||| ||||
Db   1482  GCCATCCTTTGCTGCTGGCAGATTATTGGAAGTGGTGAATCGGTGCAAAACGGTGAATGAACTTCA  1541

Qy   1171  GAATTGACGGCAGCTGAGGTGGTAGTACCAAGAGACACCCCTGATGAGAACGACCA  1230
                |||||  ||| || || ||| ||  ||| || ||  |||| ||  |||| ||||
Db   1542  GAATTTGTCAAGTGCAGAAGTTGTTGTCCCTGACACCTGACCAGAATGAGAATGACCA  1601

Qy   1231  GGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCG  1290
                |||||| || |||||| || | |||  ||  | |||| || | |  |||| |||| |||
Db   1602  AGTGGTTGTCAAAATAACTGGTCACTTCTATGCTTGCCAGGTTGCCCAGAGAAAAATTCA  1661

Qy   1291  AGACATCCTGGCCCCAGGTTAAGCAGCAGCATCAGAAG  1327
                |  ||| |  || |||||| |||| ||| || ||| ||
Db   1662  GGAAATTCTGACTTCAGGTAAAGCAGCACCAAACAACAG  1698
```

Search completed: July 13, 2004, 18:31:39
Job time : 707 secs

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:     July 13, 2004, 12:03:32 ; Search time 30 Seconds
                                     (without alignments)
                                     10953.008 Million cell updates/sec

Title:          US-09-270-437D-5
Perfect score:  3110
Sequence:       1 agggacgctgcgcaccgcc........attcctcaggtttaaaa 1708

Scoring table:  BLOSUM62
                Xgapop  10.0 , Xgapext  0.5
                Ygapop  10.0 , Ygapext  0.5
                Fgapop   6.0 , Fgapext  7.0
                Delop    6.0 , Delext   7.0

Searched:       283366 seqs, 9619526 residues

Total number of hits satisfying chosen parameters:    566732

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match  0%
                 Maximum Match  100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_13072004_121922_9618/app_query.fasta_1.1863
-DB=PIR_78 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09270437 @CGN_1_1_44 @runat_13072004_121922_9618 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   PIR_78:*
             1:  pir1:*
             2:  pir2:*
             3:  pir3:*
             4:  pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	451.5	14.5	854	2	T23837	hypothetical prote
2	265	8.5	621	2	D96554	hypothetical prote
3	256.5	8.2	398	2	T41600	probable pre-mRNA
4	251	8.1	479	2	C86275	hypothetical prote
5	248	8.0	768	2	T27855	hypothetical prote
6	244	7.8	644	2	A53184	myc far upstream e
7	240.5	7.7	680	2	T25832	hypothetical prote
8	235.5	7.6	632	2	T02627	hypothetical prote
9	234.5	7.5	568	2	T49962	hypothetical prote
10	232	7.5	589	2	T19216	hypothetical prote
11	232	7.5	611	2	T19217	hypothetical prote
12	221.5	7.1	510	2	I38489	onconeural ventral
13	218.5	7.0	356	2	S58529	alpha-complex prot
14	217.5	7.0	396	2	S41224	hnRNP protein - Af
15	214	6.9		2	T48439	probable RNA-bindi
16	209	6.7		2	T30168	hypothetical prote
17	209	6.7		2	A44125	high density lipop
18	203.5	6.5		2	S23464	vigilin - chicken
19	202.5	6.5		2	T04533	hypothetical prote
20	199	6.4		2	T04571	hnRNP protein E2 -
21	197.5	6.4		2	S42471	dC stretch-binding
22	197	6.3		2	S41495	transformation upr
23	194	6.2		2	S43363	single-stranded nu
24	194	6.2		2	S78515	rna binding protei
25	193	6.2		2	T41389	kappa-B motif-bind
26	191	6.1		2	A54143	hnRNP complex prot
27	190	6.1		2	S46109	protein C08H9.2 [i
28	190	6.1		2	B88279	hypothetical prote
29	182.5	5.9		2	T19117	SCP160 protein - y
30	178	5.7		2	S56030	hypothetical prote
31	164	5.2	c	2	S45766	hypothetical 82K p
32	163	5.2		2	JQ0317	hypothetical prote
33	159	5.1		2	T13690	F14M2.18 protein -
34	156.5	5.0		2	C86460	probable proline-r
35	155.5	5.0		2	G84693	hypothetical prote
36	155.5	5.0		2	S42933	hypothetical 24.7K
37	152	4.9		2	JQ0320	related to SCP160
38	151.5	4.9		2	T49692	hypothetical prote
39	148	4.8		2	T04255	MHC class III hist
40	146	4.7		2	S36152	hypothetical 119.5
41	144	4.7		2	JQ0405	hypothetical prote
42	144	4.6		2	T04011	MHC class III hist
43	143.5	4.6		2	S37671	nascent polypeptid
44	143	4.6		2	T30826	glutenin high mole
45	142	4.6		2	B35098	MHC class III hist

ALIGNMENTS

```
RESULT 1
T23837
T23837
hypothetical protein M88.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T23837
R;Sulston, J.
submitted to the EMBL Data Library, June 1994
A;Reference number: Z19806
A;Accession: T23837
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-854 <WIL>
A;Cross-references: EMBL:Z34802; PIDN:CAA84338.1; GSPDB:GN00021; CESP:M88.5
A;Experimental source: clone M88
C;Genetics:
A;Gene: CESP:M88.5
A;Map position: 3
A;Introns: 38/3; 96/1; 133/3; 255/3; 307/3; 395/2; 436/2; 492/3; 657/3; 759/1; 820/1
```

Alignment Scores:
```
Pred. No.:                          1.21e-24        Length:                854
Score:                              451.50          Matches:               158
Percent Similarity:                 39.38%          Conservative:          96
Best Local Similarity:              24.50%          Mismatches:            225
Query Match:                        14.52%          Indels:                166
DB:                                 2               Gaps:                  20
```

```
US-09-270-437D-5 (1-1708) x T23837 (1-854)

Qy    40 CATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAGGTCTCCTACATCCC 99
            ||||||||||::::.:::  ||||||||||||::::  ::::: ::      |||
Db   241 HisHisGlnLeu-HisProGlnMetMetGlnGlnHisAla---GlnLeuTyrHisPr 259

Qy   100 CGATGAGCAGATAGCACAG---------------GGACCTGAGAATGGGCGCCGAGGGGGCTTGG 150
            :::||||  |||                   |||
Db   259 cHisGlnGlnAsnGlnHisGlnAlaGlyGlnHisGlnHisGlnSerHisHisGlnSerGl 279
```

```
Qy    151  CTCTCGGGTCAGCCCCGCAGGGCTCACCTGTGCCAGCGGGGCCCAGCCAAGCAGCA  210
                :::||||                           ::||||||
Db    279  nAsnHisAsnGlnHisArgAsnHisArgAsnGlnSerHisSerGlyProHisHisIleProGl  299

Qy    211  GCAAGTG----------------------GACATCCCCTTCCGCTCCTGGTTGCCCAC  246
                 :::||             |||  ||:||||    ::||||
Db    299  nAsnLeuMetMetProArgCysMetLeuLysAspTrpProIleArgCysValValGluGl  319

Qy    247  CCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCAACATCACAAAACA  306
              ||||       |||||||||:::|||  :::||||||
Db    319  yLysTyrHisAlaValIleIleGlyProAsnGlySerThrIleLysAspIleAlaSerSe  339

Qy    307  GACCCAGTCCAAGATAGAC-------GTGCATAGAAGGAGGAACGCA-----GGTGC  351
                ||||  :::||||||:::||            :::||||||
Db    339  rThrArgCysLysAspPheValAsnLeuSerLysGlyArgThrLeuGlyAs  359

Qy    352  AGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCGTGCTCCTCCGCTTGTAAGAT  411
                ||||||||:::|||||||           :::||
Db    359  nAsnAspArgIleLeuThrValHisGlyValAlaGluGlnAlaThrLysAlaValAlaAr  379

Qy    412  GATCTTGGAGATTATGCATAAAGAGGCT-----AAGGACACCAAAACGGCTGACGAGT  465
             ||||||      |||||||:::|||           :::|||
Db    379  gIleLeuAspValIleGlnSerGluAlaValLysAspAspValAsnValGlyAlaAspTh  399

Qy    466  TCCCCTGAAGATCCTGCCCAATAACTTTGTAGGGCGTCCATTGGCAAGGAGGACG  525
              |||||||||:::::||||||||||||
Db    399  rValLeuArgMetArgAlaHisAsnGlnLeuCysGlyArgLeuIleGlyLysAlaGlySe  419

Qy    526  GAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCC------  574
              :::||||     |||  |||||||||||||:::::
Db    419  rSerIleLeuLysGluIleMetGlnLysThrGlyThrAsnIleThrValThrLysTyrIleg  439

Qy    575  ------TCGTTGCAAGACCTTACC----------CTTTACAACCCTGAGAGGAC  612
                |||||                      :::||
Db    439  uProProGlyGlyIleSerGlyLeuThrLeuAlaAsnGluLeuLeuLeuLeuGluMetGluArgTh  459

Qy    613  CATCACTGTGAAGGGG----GCCATCGAGAATTGTTGCAGGGCCCAGCAGCAGAAATAATGAA  669
             |||  :::||||       |||||||||||||:::||
Db    459  rIleMetValArgGlyProSerIleGluAlaValAlaValGlnAlaLeuGluAlaLeuGluSerAl  479

Qy    670  GAAAGTTCGGAGGCCTATGAGAATGATGGCTGCCATGAGCTCCATGGGCTCACCTG-------  721
             ||:::                |||||||||||:::||
Db    479  aLysLeuLysCysTyrGluSerGlnLeuuArgAlaGlnLnSerMetGlnLysPr  499

Qy    722  -ATCCCTGGCCTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGAGTCCC  780
               ::||||     ::|||   |||||||:::||
Db    499  oMetProProMetMetProProIleLeuProProGlyAlaSerSerSerAlaValSe  519

Qy    781  G----------------CCGCTCCCAGCAGCGTTACTGGGGCTGCTCCCTAT-------  817
              |                ::||||||  ::||
Db    519  rAlaProHisPheIleProThrProValGlyValMetGlnHisValGlnHisPheAlaSerSe  539

Qy    818  -----------------AGCTCCTTATGCAG----------  832
                         :::||||:::||
Db    539  rGlnHisLeuValHisGlnAsnAlaAsnSerPheLeuGlnProGlyValLeuGlnIl  559

Qy    833  ------------GCTCCCGAGCAGGAGATGTGCAGGTGTTTATCCCGCCCCAGGCCAGTGGG  882
                      :::||||:::|||||||
Db    559  eGlnProGlyThrThrAsnLeuArgGlnValArgMetTrpProValProAspSerMetIleGl  579

Qy    883  CGCCATCATCCGGCAAGGAGGGGCCAGCACACATCAAACAGCTTCCCCGGTTTGCCAGCGCCTC  942
             |||||||||:::||||     ||:::||||
Db    579  yAlaLeuIleGlyAlaLysGlyLysAsnIleLysMetIleIleAArgAspThrGlyAlaSe  599

Qy    943  CATCAAGATTGCACCACCCGAAACACCTGACTCCAAA------------------  979
             |:::||||||||                   :::
Db    599  rValLysIleGluAlaProGluGluLysThrGlnArgGluLeuLysThrGluGluLysLy  619

Qy    979  ------  979
Db    619  sArgLysLeuAspThrAspSerGlyGluCysValValAlaSerGlyHisAspProGl  639

Qy    979  ------  979

Db    639  nGluPheLeuGluAspAsnAlaThrIleAsnSerSerAspAlaIleGluGluLysProGly  659
Qy   1030  GGGAAGAATCTATTGGCAAACTCAAGGAGGAGGAGAACTTCTTTGGTCC--------  1075
Db    659  sProValSerGluArgMetValThrIleAsnGlyAspAspLeuGlnLeuLeuLysAlaGl  679
             :::||
Qy   1076  ------AAGGAGGAAGTGAAGCTGGAGACCCCACATACAGTCGTGCCAGCATCAGCAGC  1125
Db    679  nSerTyrValPheSerLysIleAlaGluThrSerSerLeuProSerSerLeuProSerGlyMetAs  699
             :::

Qy   1126  TGGCCGGGTCATTGGCAAAGGTGCAAAAAACGGTGAAAAACGGTGAACCAGTTGCAGAATTGACGGCGAGC  1185
             |||||||::::||
Db    699  pGlyAspArgSerHisMetLeuArgIleArgThrGluLysArgThrGluValSerProThrArgIleIl  719

Qy   1186  TGAGGTGGTAGTACCAAGAGACCAGACCCCTGAT----------------  1219
             |||||||:::||
Db    719  eGlyArgIleIleGlyLysGlyGlyGluAsnValArgGlnLeuArgIleThrGlyAl  739

Qy   1220  ---GAGAACGACCAGGTCATCGTGAAAATCATCGGACAATTCTATGCCAGTCAGATGGC  1275
             :::||||
Db    739  aValValLysIleProArgGluGluGlyLeuArgAsnGlyGlyGluValTyrArgHisAspAspGl  759

Qy   1276  TCAACGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAGGGGACAGAG  1335
             |||||||||:::||
Db    759  yLeuGluGluArgMetMetIleArgThrIleGlyAsnMetTyrSerThrHisAsnVa  779

Qy   1336  TAACCAGGCACGGAAGTGAACGACCACCAGCCCCTCCCTGTCCCTTNGAGTCCAGGAC  1395
             :::||||       ||||      :::
Db    779  lGlnPheArgLeuAlaHisLeu------ValAsnGluTyrTyrArgSerGlyAspHi  796

Qy   1396  AACAACGGGCAGAAATCAGCAGTGTGCTCTCCCCGGCAGGCCTGAGAATGAGTGGGAATC  1455
                                     :::||||
Db    796  sArgAsnLysSerSerAspTyrLys  804

Qy   1456  CGGGACACNTGGGCCGGGTGTGAGATCAGGTTTGCCCACTTGATTGAGAAGATGTTCCA  1515
             :::                |||||:::
Db    805  ------GlyGlyArgProHisSerAlaProSerSe  814

Qy   1516  GTGAGGAACCCTGATCTNTCAGCCCCAAACACCCCAATTGGCCAACACTGTNTGCC  1575
             |||:::                          ||||||
Db    814  -----------------LeuGluLysMetAspGl  828

Qy   1576  CCTCGGGGTGTCA  1588
             :::||||
Db    828  nLeuGlyThrIle------AlaProIleSerArgAsnSerArgAlaSerPr  843

Qy   1589  oLysSerValSer  847
```

RESULT 2
D96554
hypothetical protein F19C24.19 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: D96554
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.R.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96554
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-621 <STO>
A;Cross-references: GB:AE005173; NID:g11094762; PIDN:AAG29695.1; GSPDB:GN00141

C;Genetics:
A;Gene: F19C24.19
A;Map position: 1

Alignment Scores:
Pred. No.:   3.65e-11
Score: 265.00
Percent Similarity: 38.66%
Best Local Similarity: 23.68%
Query Match: 8.52%
DB: 2

US-09-270-437D-5 (1-1708) x D96554 (1-621)

```
Qy  221 ATCCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGG 280
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   19 ValHisPheArgLeuLeuValProAlaThrArgThrGlyAlaIleIleGlyLysGlyGly 38

Qy  281 GCCACCATCCGCAAACTACACAAACAGACACCAGTCCAAGATA---------------- 322
        :::||||||:::::::::    |||||||
Db   39 SerValIleArgHisLeuGlnSerValThrGlySerLysIleArgValIleAspAspIle 58

Qy  322 ------------------------------------------------------------ 322
Db   59 ProValProSerGluGluArgValValLeuLeuIleIleAlaProSerGlyLysLysAsp 78

Qy  323 GACGTGCATAGGAGGAGAACGCAGGTGCAGCTGAA------------------------- 358
        :::||       ||||||     ||||||
Db   79 GluSerAsnValCysAspSerGluAsnProGlySerGluGluProLysGlnGluLysGly 98

Qy  359 -----AAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTGTAAG 409
             ||||||||||      |||||||
Db   99 SerGluCysAlaGlyThrSerGlyGlyAspAspGluGluAlaProSerSerAlaGlnMet 118

Qy  410 ATGATCTTGGAGATTATGCATAAA----------GAGGCTAAGACACCAAAAGCGCTGAC--- 460
        :::          |||       ||| :::
Db  119 AlaLeuLeuArgArgValPheGluValPheGlyAspAspAspAlaAlaThrValAspGly 138

Qy  461 --------------------GAGGTTCCCGTGAAGATCCTGCCCATAATAACTTT 496
                            :::||||||
Db  139 AspGluLeuAspLysGlyLeuSerGluGluLysLeuCysArgMetIleValArgGlyAsnGln 158

Qy  497 GTAGGGCGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAGAAGGTAGAGCAAGATACCGAG 556
        |||::: |||||     :::   ::||||:::
Db  159 ValAspTyrLeuMetSerLysGlyLysGlyGlyLysMetIleGlnIleLysIleArgGluAspSerGly 178

Qy  557 ACAAAAATCACCATCTCCTCGTTGCAA------------------------GACCTT 589
        :::   |||||||      |||
Db  179 AlaIleValArgIleSerSerThrAspGlnIleProProCysAlaPheProGlyAspVal 198

Qy  590 ACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGGCC-----------ATCGGAAT 640
        :::  :::|||||      |||
Db  199 ValIleGlnMetAsnGlyLysPheSerSerValLysSerAlaLeuLeuLeuValThrAsn 218

Qy  641 TGT------------------------------TGCAGGGCCGAGCAGGAGAAATA 664
        |||
Db  219 CysLeuGlnGluSerGlyAlaProThrTrpAspGluCysProPheProGln-------------236

Qy  665 ATGAAGAAAAGTTCGGGAGGCCTATGAGATGATGTGGCTGCCATGAGCTCTCACCTGATC 724
                                    :: :::
Db  237 -----------------ProGlyTyrProGluTyrHisSerMetGluTyrHis--------249

Qy  725 CCTGGCCCTGAACCTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG 784
        |||:::|||                    ||||||
Db  250 -----------------ProGlnTrpAspHis-------ProPro-----------256

Qy  785 CCTCCCCAGCAGCGTTACTGGGGCTGCTCCCCATAGCTCCTTTATGCAGGCTCCCGAGCAG 844
        |||:::          :::||||||
Db  257 ProAsnProMetProGluProProPheAsnArgProVal---ValGluGluGlu 275

Qy  845 GAGAGTGTGCAGGTGTTTATCCCGGCCACGAGTGGGCCATCATCGGCAAGGAGGGG 904
        |||:::  |||||||||:::||||||
Db  276 ValAlaPheArgLeuLeuCysProAlaAspLysLysProMetAspMetAspAla 295
```

Alignment Scores:
Pred. No.:   1024
Score: 334
Percent Similarity:
Best Local Similarity:
Query Match: 1066
DB: 354

```
Qy  905 CAGCACATCAAACAGCTCTCCCGGTTGCCAGCGGCTCCATCAGATTGACACCCGAA 964
        :::||||             ||| :::
Db  296 AlaValAlaArgAlaLeuGlnAsnGlnAsnGlySerGlyAlaSerIleLeuLysValSerAlaSerArgAspPro---314

Qy  965 ACACCTGACTCCAAAGTTGTATGGTTATCATCACTGGACACGCCAGAGGCCCAATTCAAG 1024
        ||| |||||||||||||||||::::::::|||
Db  315 ThrHisAspSerGluGluArgValIleLeuValIleLeuSerAlaIleArgAspGluArgArg 334

Qy 1025 --------GCTCAGGGA----------AGAATCATGGCAAACTCAAGGAGGAGAACTTC 1066
                |||||||||         |||||||
Db  335 HisSerLeuAlaGlnAsnAspGlyValMetArgValHisAsnArgIleValGluIleGlyPhe 354

Qy 1067 TTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCT 1126
        :::        |||||||
Db  355 --------GluProSerAlaAlaValAlaValAlaArgLeuLeuLeuValHisSerProTyrIle 371

Qy 1127 GGCCCGGTCATTGGCAAAGTGGAAAAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCT 1186
        |||||||||::::::::|||::::||||||
Db  372 GlyArgLeuLeuGlyLysGlyGlyValHisLeuIleSerGluMetArgArgAlaThrGlyAla 391

Qy 1187 GAGGTG---GTAGTACCAAGAGACCAGACCCCTGATGAGACGACCAG------GTCATC 1237
        :::       |||       |||
Db  392 SerIleGluArgValPheAlaLeuAlaLysGlnAlaThrLysTyrGluSerGlnHisAspGluIle 411

Qy 1238 GTGAAAATCATCGGACATTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATC 1297
        :::||||:::::|||
Db  412 ValGlnIleValGlyAsnLeu--------LysThrValGlnAspAlaLeuPheGlnIle 428

Qy 1298 CTGGCCCAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAACCAGCCCAGGCCAGCGGAGG 1357
        |||::: |||||||
Db  429 LeuCysArgLeuArgGlyAlaMetPheProGly--ArgLeuProPheGlnGlyMetGlyGl 448

Qy 1358 AAGTGACCAGCCCTC-----------------------CCTGTCCCTTGAGTCCAGGACA 1396
        :::|||
Db  448 yProProProMetGlyProTyrProGluProProPheGlyProProArgGl 468

Qy 1397 ACAACGGGCAGAAATCGAGAGTGTGCTCTCCCCGGCAGGC 1436
        |||
Db  468 nTyrProAlaSerProValGly 481
```

RESULT 3
T41600
probable pre-mRNA splicing factor - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T41600
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z22003
A;Accession: T41600
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-398 <LYN>
A;Cross-references: EMBL:AL031825; PIDN:CAA21234.1; GSPDB:GN00068; SPDB:SPCC757.09c
A;Experimental source: strain 972h-; cosmid c757
C;Genetics:
A;Gene: SPDB:SPCC757.09c
A;Map position: 3

Alignment Scores:
Pred. No.:   1.34e-10
Score: 256.50
Percent Similarity: 44.74%
Best Local Similarity: 27.19%
Query Match: 8.25%
DB: 3

Length: 398
Matches: 93
Conservative: 60
Mismatches: 130
Indels: 59
Gaps: 14

US-09-270-437D-5 (1-1708) x T41600 (1-398)

```
Qy  161 CAGGGCCCGC------------CAGGGCTTCACCTGGCGAGCGGGGCCCA---------199
        ||||||                    |||:::
Db   68 GlnProGluProThrSerGlnValProIleSerAlaLysProMetAspAspAla 87
```

Length: 621
Matches: 117
Conservative: 74
Mismatches: 184
Indels: 119
Gaps: 19

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C;Accession: C86275
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86275
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-479 <STO>
A;Cross-references: GB:AE005172; NID:g5080792; PIDN:AAD39302.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Alignment Scores:
Pred. No.:   Length: 479
Score: $3.52e{-10}$ Matches: 100
Percent Similarity: 251.00   Conservative: 89
Best Local Similarity: 44.89%   Mismatches: 140
Query Match: 23.75%   Indels: 92
DB: 8.07%   Gaps: 19

US-09-270-437D-5 (1-1708) x C86275 (1-479)

[Nucleotide and protein sequence alignment rows Qy/Db continue]

RESULT 4
C86275
hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)

```
Db   228  HisLeu----------------------------------LeuLeuSerSerSerSer  237
Qy   776  GTCCCCGCGCTCCCAGCAGCGTTACTGGGCTGCTGTCTCCCTATAGCTCCTTTATGCAG---  832
                      ||||||  :::    :::     ||||||   :::::::
Db   238  MetHisGlnProGlyAlaMetLeuMetSerAlaAlaLeuThrSerSerHisArgAsnTyr  257
Qy   833  -----------------GCTCCCAGCAGGAGATGTGCAGGTGTTATC---CCCGCC  871
                           ::||||         ||||||   ||||||
Db   258  AlaValArgArgAspIleAlaAspAlaArgGlyPheGluPheCysValCysPheIleCysProAla  277
Qy   872  CAGGCAGTGGCGGCCGCCATCATCGGCAAGAAGGGGCAGCACATCAAAACAGTCTCCCGGTT  931
                      :::   :::||||||   :::     :::
Db   278  GluAsnValGlyValIleGlyLysGlyValGlyPheIleAsnGlnIleArgGlnGlu  297
Qy   932  GCCAGCGGCCTCCATCAAGATTGCACCCGAAACACCTGACTCCAAGTTCGTATGGTT  991
                      ||||||  :|||::::::   ||    :::::::
Db   298  ThrGlyAlaThrIleArgValAsnThrSerGluThrAspAspAspAsp------CysIle  315
Qy   992  ATCATCACTGGACCGCCAGAG--------------------------GCCCAATTC  1021
                      |||
Db   316  IlePheIleSerLysGluPheTyrGluAspGlnSerProAlaValAsnAlaAlaIle  335
Qy  1022  AAGGCTTCAGGGAAGAAGAATCTATGGCAAACTC---AAGGAGGAGAACTTCTTTGGTCCCAAG  1078
                      :::     ||||||         :::||||||
Db   336  ArgLeuGlnGlnArgCysSerSerGluLysValGlyValGlyLysLysAspAlaAsn---------  350
Qy  1079  GAGGAAGGTGAAGCTGGAGACCCACATACGTGTCCAGCATCAGCAGCTGGCGCGGGTCATT  1138
                      :::     ||||  :::||||||  ::::||
Db   351  ---AspLeuAlaIleSerThrArgLeuLeuValSerSerSerGlnIleGlyCysLeuIle  369
Qy  1139  GGCAAAGGTGAAAAACGGTGAACGAGTTGCAGATTTGCACGGCCAGCTGAGGTG---GTA  1195
                      |||||||||||      |||   :::   ::::
Db   370  GlyLysGlyGlyLysAlaValIleSerGluMetArgSerValThrArgAlaAsnIleArgIle  389
Qy  1196  GTACCAAGAGACCAGACCCCTGAT------GAGACGACCAGGTCATCGTGAAAATCATC  1249
                      ||||||||||  :::::::::::||:::
Db   390  LeuGlnLysGluAspValProLysIleAlaArgGluAspGluMetValGlnIleThr  409
Qy  1250  GGACAATTCTATGCCAGTCAGATGGCTCAACGGAGGCTCCAGAGATCCTGGCCCAGGTT  1309
                      |||      :::         :::   ::::::
Db   410  Gly---------SerProAspAlaAlaMetLysAlaMetLysAlaLeuThrGlnValIleIleLeuArgLeu  426
Qy  1310  AAG  1312
               :::
Db   427  Arg  427
```

```
US-09-270-437D-5 (1-1708) x T27855 (1-768)

Qy   131  GGGGGCCGAGGGGGCTTTGGCTCTGGGCTGGCGTCAGCCCCGCCAGGGCTCACCTGTGCAGCG  190
                   ||||||   :::     :::
Db     1  GlyAlaAspGlyGlyTyrTrpLeuGlyTyrLeuLysArgProLeuAsp---SerGluIleLeuAsp  242
Qy   191  GGG-----GCCCCAGCCAAGCAGCAGCAA-----------------  214
               |||
Db   243  GlyAspLeuIleProThrLysLysSerSerGluValGlyValGlyAspLeuAsnMetGlyAspSer  262
Qy   215  --------GTGGACATCCCCCTTCGGCTCTGGTGCCCACCCAGTATGTGGGTGCCATT  265
                   :::       :::
Db   263  AspLysIleThrAspIle--------TyrProValProGluLysValValGlyLeuVal  279
Qy   266  ATTGGCAAGGAGGGGGCCCACCATCCGCAACATCAAAACAGACCAGTCAAGATAGAC  325
                   ||||||   :::     :::     :::::::
Db   280  IleGlyLysGlyGlySerGluIleArgLeuLeuGlnGlnThrSerGlyValCysArgValGln  299
Qy   326  GTGCATAGGAAGGAGAGCCAGGTGCAGTGAAAAGCCATCAGTGTGCACTCCACCCCT  385
                   ::      ||||||   :::     :::
Db   300  MetAspProAsnHisGlnSerValAsnGlyPheArgAsnCysThrIleGluGlyProPro  319
Qy   386  GAGGGCTGCTCCTCCGCTGTAAGATGATCTTGGAGATTATGCATAAAGAG-------  436
                   ::||||||       ::::||||
Db   320  AspGlnValAlaValAlaArgGlnMetIleThrGlnValIleAsnArgAsnGlnThrGly  339
Qy   437  GCTAAGGACACCAAAAACGGCTGACCGAGTTCCCCTGAAGATCCTGGCCCATAATAACTTT  496
                   ||:::       ||||||   :::
Db   340  AlaGlnProGlyAlaAlaProGlyValThrGluValThrArgLeuIleProAlaAspLys  359
Qy   497  GTAGGGCGTCTCATTGGCAAGGAAGGACGGACGAACCTGAGAGGCAAGAAGTAGACCGAG  556
                   ::||||||:::   :::||:::
Db   360  IleGlyLeuValIleGlyLysGlyGlyGluThrIleLysGlyLeuGlnGluGlnSerGly  379
Qy   557  ACAAAAATCACCACCTCCTGTTGCAAGACCTTACCCTTTAC-------AACCCT  604
                   :::     ::||:::   |||
Db   380  LeuArg----AsnCysAsnValGlnGluThrProThrArgAlaThrArgGlyGlnProLysPro  398
Qy   605  GAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTGCAGGGCCAGCAGGAAATA  664
                   |||     ::::   ::::::
Db   399  LeuArgMetIleGlySerProAlaAlaIleAlaGluThrAlaLeuAlaLeuValHisAsnIle  418
Qy   665  ATGAGAAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGCTCCATGAGCTCTCACCTGATC  724
                   ||:::
Db   419  MetAsnAsnThrGln  423

US-09-270-437D-5 (1-1708) x T27855 (1-768)

Qy   725  CCTGGCCTGAACCTGCGTCGTGTAGTCTTTTCCCAGCTTCATCCAGCGGCAGTCCGCGCCG  784
                   |||         :::
Db   424  --------------------GlyAsnAlaProLeuGluLeuGlnArgAla---ProHis  434
Qy   785  CCTCCCAGCAGTTACTGGGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCAGCAGCAGCGGG---  844
                   ||||||       ||||||
Db   435  GlnProSerGlyGlnPheGlyGlyTyrGlyTyrGlyAlaGlnLeuAlaGlnAlaLysGlyGly--  453
Qy   845  GAGATGGTGCAGGTGTTTATCCCCCGCCAGGCAGTGGGGCCCATCATCGGCAAGAAGAGGGG  904
                   :::||:::   :::
Db   454  --------GluValIleValProArgLeuSerAlaGlyMetIleIleGlyLysLysGlyGly  470
Qy   905  CAGCACATCAAAACAGCTCCCGGTTGCCAGCGGCCTCCATCAAGATTGCACCCCCGAA  964
                   ||||:::||   :::||     :::
Db   471  GluMetIleLysArgLeuAlaLaGluThrGlyThrLysIleGlnPheLysProAspThr  490
Qy   965  ACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCCCAGAGGCCAATTCAAG  1024
                   ||:::||||||   :::     :::::::
Db   491  AsnProAsnSerGluAsnArgIleAlaValIleIleMetGlyThrArgAspGlnIleTyrArg  510
Qy  1025  GCTCAGGGAAGAATCATGGCAAACTCAAGGAG-----GAGAACTTCTTTGGTCCCAAG  1078
                   ||||||   :::||   ::::||
Db   511  AlaThrGluArgIleThrGlnValIleGluValAlaAsnArgAlaIleLeysAsnGlyAlaProGln  530
Qy  1079  GAGGAAGTGAAGCTGGAGACC---------------------CACATACGTGTGCCA  1114
               :::::::
```

```
RESULT 5
T27855
hypothetical protein ZK418.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27855
R;Fulton, L.
  submitted to the EMBL Data Library, April 1994
A;Description: The sequence of C. elegans cosmid ZK418.
A;Reference number: Z20430
A;Accession: T27855
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-768 <FUL>
A;Cross-references: EMBL:U00047; PIDN:AAA50693.1; CESP:ZK418.9
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:ZK418.9
A;Introns: 166/1; 221/3; 243/3; 279/1; 387/3; 490/1; 567/3; 727/2

Alignment Scores:
Pred. No.:
Score:                  6.59e-10           Length:        768
                        248.00             Matches:       131
Percent Similarity:     37.89%             Conservative:   88
Best Local Similarity:  22.66%             Mismatches:    225
Query Match:            7.97%              Indels:        134
DB:                     2                  Gaps:           22
```

Db   531 AspArgGlySerAlaGlyThrValLeuProGlyGlnSerIlePheTyrMetHisValPro 550

Qy   71 AACCATGCCTGAAGGTCTTCCTACATCCCGATGAGCAGATGAGACAG------GGACCT 124
                ||| |||||||||::::||||| ||||||
Db   1115 GCATCAGCAGCTGGCCCGGTCATTGGCAAAGTGGAAAAACGGTGAAGAGTTGCAGAAT 1174

Qy   28 AsnAspAlaPheLysGlyLysGlyLeuGlnIleAlaArgGlnIleAlaLysIleGlyGly 47
Db   551 AlaGlyLeuLysCysGlyValIleGlyLeuValIleGlyGlyGluAsnIleLysGlnIleGluArg 570

Qy   125 GAGAATGGG--------CGCCGAGGGGGCTTGGCTCTGCGGGTCAGCCCCCAGGGC 175
                :::|||| ::::||||| |||||||||||
Db   1175 TTGACGGCAGCTGAGTG---GTAGTACCAAGAGACCAGAGAACGACCAG 1231

Qy   48 AspAlaGlyThrSerLeuAsnSerAsnAspTyrGlyTyrGlyGlyGlnLysArg------ 65
Db   571 GluThrGlyAlaThrCysGlyLeuAlaProAlaAlaGluGlnLysAsnGlu--AspGlu 589

Qy   176 TCACCTGTGGCAGCGGG--------------------------------- 193
                |||
Db   1232 GTCATCGTGAAAATCATGGACAT-----TTCTATGCCAGTCAGATGGCTCAA 1279

Qy   66 ---ProLeuGluAspGlyAspGlnProAspAlaLysLysValAlaProGlnAsnAspSer 84
Db   590 LysValPheGluIleLysGlySerGlnLeuGlnIleHisHisAlaSerHisLeuValArg 609

Qy   194 --------GCCCCAGCAAGCAGCAGCAGCAAGTGGACATCCCCTT------CGG 232
                        |||||
Db   1280 CGGAAGATCCGAGACATC-------------------- 1297

Qy   85 PheGlyThrGlnLeuProPheMetHisGlnGlnGlnGlnGlnSerArgSerValMetThrGluGlu 104
Db   610 IleLysLysValGlyGluIleuSerProAsnThrProValProProLeuGlnGlyAlaGlyGly 629

Qy   233 CTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATCCGC 292
                                            ||||||
Db   1297 -------------------------------------- 1297

Qy   105 TyrLysValProAspGlyMetValGlyPheIleIleGluGlyGlyGluGlnIleSer 124
Db   630 GlyTyrGlnGlnGlnGlnAlaMetPheSerAlaGlyThrGlnAsnGlyGlyTyrGlnSer 649

Qy   293 AACATCACAAAACAGACCCAGTCCAAGATAGACGTTCAAGAAGGAGAACGCAGGTGCA 352
                                                CAGCATCAGAAGGACGACAGAGT
Db   1298 ------------------CTGGCCCAGTTAAGCAG 1336

Qy   125 ArgIleGlnGlnSerGlyCysLysLysValGlnIle--AlaProAspSerGlyValGlyLeu 143
Db   650 ThrGlyGlyGlyPheIleGlnGlnGlnGlnGlnGlnProGlnPheGlnGlnGlnGlnGlnTrp 669

Qy   353 GCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGCTGCTCCTCCGCTTGTAAGATG 412
                |||| |||||
Db   1337 AACCAGGCCCAGGCACGGACGACACGGAGAAGTGACCAGCAGCAAAATCACCATCTCCTCG 1396

Qy   144 ProGluArgSerCysMetLeuThrGlyThrProGluSerValGlnSerAlaLysArgLeu 163
Db   670 AlaProGlnAsnGly--GlyGlyValGlnIleArgLeuProThrGluMetTyrGluMetPr 689

Qy   413 ATCTTGGAGATTATGCATAAA------GAGGCTAAGGACCACCAAAACGGCTGAC---- 460
                ::: |||||||||
Db   1397 ACAACGGGCAGAAATCGAGAGTGTGTCTC-----TCCCCGCAGG 1435

Qy   164 LeuAspGlnIleValGluLysGlyArgProGlyPheHisHisGlyAspProGlyPro 183
Db   689 oMetGlnGlnProGlnAlaSerValIleProGlnAlaSerAlaAspProAlaProAlaVa 709

Qy   461 --GAGGTTCCCTGAAGATCCTGGCCCCAATAATAACTTGTAGGGCGTCTCATTGGCAAG 517
                                   ||||||
Db   1436 CCTGAGAATGAGTGGGAATCCGGGAC-------ACNTGGGCCCGGCTGTGTAGATCCAGGT 1486

Qy   184 GlyAsnAlaValGlnGluIleMetIleProAlaSerLysAlaGlyLeuValIleGluGlyLys 203
Db   709 lAsnProAlaThrGlyAlaPro-AspTyrSerAlaGlnInTrpAlaAlaTyr---TyrArgS 728

Qy   518 GAAGGACCGGAACCTGAAGAAGTAGAGCAAGATACCGAGACAAAATCACCATCTCCTCG 577
                                ||||
Db   1487 TTGCCCACTTGATTGAAGAAGATGTTCCAGTG-----AGGAACCCTGATC 1531

Qy   204 GlyGlyGlyGluThrIleLysGlnLeuLysGlnLeuArgGlnAlaGlyValLysMetValMet- 221
Db   728 erIleGlyLeuValGluGlnIleAlaAlaMetValGluSerGlnMetArgArgAsnGlnSer- 747

Qy   578 TTGCAAGACTTACCCTTTACAACCCTGAGACGACC-----ATCACTGTG 622
                |||||||| :::
Db   1532 TNTCAGCCCCAAACACCACCAATTGGCCCAACACTGTNTGCCCCTCCG 1581

Qy   222 IleGlnAsp---------GlyProGlnAsnThrGlyAlaAspLysProLeuAlaArgIle 237
Db   748 ------AsnThrAlaProLeuGlyAlaProLeuGlyAspValSerGlyValProGlyGln 760

Qy   623 AAGGGGCCATCAGAGAATTGTTGCAGGGCCGAGCAGAGAAATAATGAAGAAAGTTCGGGAG 682
                                ||||||::::||||::|||||| ::||||
Db   1623 ThrGlyAspProTyrLysValGlnAlaAlaLysGluMetValLeuGluLeuIleArgAsp 257

Qy   683 --------GCCTATGAGAATGATGTGGCTGCCATGAGCTTCACCTGATCCCT 727
                          ||||||:::
Db   683 GlnGlyGlyPheArgGluValArgGluValTyrGly::: 258

Qy   728 GGGCTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGAGTCCCGGCGCCT 787
                ||||||||||||::::|| |:||||||||||||||||||| |||||
Db   728 CACATCAAACAGCTCTCCGGTTTGCCAGCGCTTCCATCAAGATTGCCAACCACGAA--- 269

Qy   788 CCCAGCAGCGTTACTGGGGCGTCTCCTCCTATAGCTCCTTTATCGAGTGGGCCAGCAGGAG 847
                ::: |||
Db   269 -------------------------------------- 269

Qy   848 ATGGTGGCAGTGTTTATCCCGCCAGGCAGTGGGGGCCATCATCGCAAGAAGGGGCAG--- 907
                ::|||||::::|||||||||||||||||||||||||||||||||||| |||
Db   270 ---SerArgIleGlyGly-------------------AsnGlu 276

Qy   908 CACATCAAACAGCTCTCCGGTTTGCCAGCGCTTGCCATCAAGATTGCCAACCACGAA--- 964
                ||||::::::|| |||
Db   297 MetIleLysValIleGlnAsnAspAlaGlyArgIleGlnPheLysProLysProGly 316

Qy   965 ---ACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACTGCGACCGGCCAGAG--GCCCAA 1018
                    ||||||:::: |||||:::
Db   317 ThrThrProGlu-----------ArgIleAlaGlnIleThrGlyProProAspArgCysGlnGln 333

RESULT 6
A53184
myc far upstream element-binding protein - human
N;Alternate names: FUSE-binding protein
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C;Accession: A53184
R;Duncan, R.; Bazar, L.; Michelotti, G.; Tomonaga, T.; Krutzsch, H.; Avigan, M.; Levens,
Genes Dev. 8, 465-480, 1994
A;Title: A sequence-specific, single-strand binding protein activates the far upstream e
A;Reference number: A53184; MUID:94170991; PMID:8125259
A;Accession: A53184
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-644 <DUN>
A;Cross-references: GB:U05040; NID:g460151; PIDN:AAA17976.1; PIDN:AAA17976.1; PID:g460152
C;Keywords: DNA binding

```
A;Residues: 1-680 <SCH>
A;Cross-references: EMBL:U88174; PIDN:AAB42272.1; GSPDB:GN00019; CESP:M01A10.1
A;Experimental source: strain Bristol N2; clone M01A10
C;Genetics:
A;Gene: CESP:M01A10.1
A;Map position: 1
A;Introns: 63/2; 97/3; 121/1; 160/1; 269/3; 411/3; 512/1; 649/2

Alignment Scores:
Pred. No.:                   1                        Length:              680
Score:                  2.23e-09                      Matches:              82
Percent Similarity:     240.50                        Conservative:         82
Best Local Similarity:  41.21%                        Mismatches:          159
Query Match:            20.60%                         Indels:               75
DB:                     7.73%                          Gaps:                 12
                        2

US-09-270-437D-5 (1-1708) x T25832 (1-680)

Qy    104 GAGCAGATAGCACAGGACCTGAGAATGGGCGCCGAGGGGGCTTTGGCTTCTCGGGGTCAG 163
              ::: |||::::       |||    |||:::          |||
Db     97 LysHisIleSerSerSerThrAspThrAlaArgGlyArgSerValGlnAlaGlyAsnProGlyGly 115

Qy    164 CCCCGCCAGGGCTCACCT------GTGGCAGGCGGGGCCCAGCCAAGCAGCAGCAGTG 217
              :::|||||||||            ||       :::
Db    116 ------GluGlySerAspGluTyrGluGluTyrGluTyrAlaProProCysLysLeuThrLysGly 133

Qy    218 GACATCCCCCTTCG--------CCTGTCGTGCCC 244
              ||||||                |||::||||Pro
Db    134 AspIleAspTyrArgValAspThrSerThrValIleLysAlaSerValSerIlePro 153

Qy    245 ACCCAGTATGTGGGTGCCATTATGGCAAGGAGGGGCCACCATCCGCAACATCACACAAA 304
              |||:::          |||:::||||:::|||
Db    154 GluGluSerValGlyLeuValIleGlyArgAsnGlyValGluIleIleSerGln 173

Qy    305 CAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAAAAGCC 364
              :::           :::         |||:::       :::
Db    174 LysSerGlyCysArgValGlnIle--ValAlaGluProSerThrGlyTyrArgSer 192

Qy    365 ATCAGTGTGCACTCCACCCTGAGGGCTGCTCCTCCGGTTGTAAGATGATCTTGGAGATT 424
              :::      :::::                  |||||       |||:::
Db    193 ValAspIleTyrGlyIleSerGluAsnIleGluValAlaLysAlaLysAsnGluVal 212

Qy    425 ATGCATAAAGAGGCTAAG----------------------TyrAla 442
              ||||||                                  |||
Db    213 ValAlaAlaArgGlyArgTyrLeuSerGlnGluProLeuProCysSerValProGlnPheGln 232

Qy    443 --------GACACCAAAAACGGTGACGAGGTTCCCCTGAAGATCCTGGCCATAATAAC 493
                      ||||         |||:::        ||||
Db    233 ProIleProAlaValSerSerAsnSerSerLysValThrIleIleIleProIleProAlaAsn 252

Qy    494 TTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAGGTAGAGCAAGATACC 553
              :::||||||||||          :::::                |||:::
Db    253 LysCysGlyValAlaIleIleGlyLysLysGlyGluGlnMetArgLysArgSerTrpThr 272

Qy    554 GAGACAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACC 613
              :::|||                 ||||::::           ||||||
Db    273 AsnCysAspPheIleLeu------IleGlnGluAsnAsnIleAlaAlaSerValLysPro 290

Qy    614 ATCACTGTGAAGGGGCCATCGACCATCCGAGAATTGTTGCAGGGCCCAGCAGGAAATGAGAGAAA 673
              :::                :::          |||:::      ||||||
Db    291 LeuGlnIleThrGlyGln--------------------------------ProLysGlu 299

Db    300 ValGluHisAla------------LysAlaLeuValAlaAspIleLeuAspGlyPhe 314

Db    315 AspGluCys---------------------------ProProAla 320

Qy    794 AGCGTTACTGGGGTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTG 853
              :::    |||    :::::                        :::
Db    321 GlyMetAlaGlyAlaAsnSerProValAlaAlaMet------------SerLeu 333

Qy   1019 TTCAAGGCTCAG------------------------------------------ 1030
              |||:::
Db    334 HisAlaAlaGluIleIleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGly 353

Qy   1031 -------GGAAGAATCTATGCAAACTCAAGGAGGAGGAGAACTTCTTTGGTCCCAAG 1078
                     ||||||              :::::
Db    354 ProGlyProGlyGlyTyrGlyArgGlyArgGlyArglyGlnGlyAsnTrpAsnMetGlyProPro 373

Qy   1079 GAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCCGGGTCATT 1138
              ::::::   |||    |||::::|||
Db    374 GlyGlyLeuGln--GluPheAsnPheIleValProThrGlyLysThrGlyLeuIleIle 392

Qy   1139 GGCAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTGACGGCAGCTGAGGTGGTAGTA 1198
              |||||||||::|||:::       :::           |||   :::
Db    393 GlyLysGlyGlyGluThrValAsnGluLeuGlnThrIleSerGlnGlnSerGlyAlaAlaArgIleAlaArgIleGluGluLeu 412

Qy   1199 CCAAGAGACCAGACCCCTGATGAGAACGACCAGGTC---ATCGTGAAAATCATCGGACAT 1255
              :::                 :::             :::
Db    413 GlnIleArgAsnProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr 432

Db   1256 -------TTCTACGTCAGTCAGATGGCTCAACGGAAGAT------CCGAGACAT 1296
                     |||         :::||||               |||
Db    433 ProGlnIleAspTyrArgAlaArgGlnLeuLlleGluGluLlysIleGlyProValAsn 452

Qy   1297 CCT--GGCCCAGGTTAAGCACGCAGCAGATCAGAGAGGGACAGAGTAACCAGGCCCAGGCACG 1353
              |||   |||||||                |||||||       |||
Db    453 ProLeuGlyProProValProHisGlyProHisGlyVal------ProGlyPro--HisG 470

Qy   1354 GA-------GGAAGTGACCAGCCCTC---------CCTGTCCCTTNGA 1386
              ||              |||:::                    :::
Db    470 lyProProGlyProGlyProGlyProMetGlyProTyrAsnProAlaProTyrA 490

Qy   1387 GTCCAGGA-------------CAACAACGGGCAGAAATCGAGAGTGTGTCTCTCCCCGG 1431
              ::|||||||
Db    490 snProGlyProProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyT 510

Qy   1395 -----------------------------------------------------------ProA 526
              :::
Db    510 rpGlyAsnAlaTyrProHisTrpGlnGlnGlnAlaProAspAsp------------ 526

Db   1432 CAGGGCCTGAGAATGGGAGTCCGGGAATCCGGGACACNTGGGCCGGGCTGAGATCAGGTTGCC 1491
              |||:::::               :::
Db    526 laLysAlaGlyAlaThrArgAspProAsnSerAla-AlaTrpAlaAlaTyr-------TyrAla 542

Db   1492 CACTTGATTGAGAAAGATGTTCCAGTGAGGGAACCCTGATCNTNCAGCCCCAAACACCCAC 1551
              :::|||                   :::::          ||||||
Db    543 HisTyrTyrGlnGlnGlnAla--------------------------GlnProProAlaAla 555

Db   1552 CCAATTGGC--CCAACACTGTNTGCCCTCGGGTGTCAGAAATTNTAGCGCAGGCAC 1608
              :::|||        :::
Db    556 ProAlaGlyAlaProThrThrThrGlnThrArgAsnGlyGlnGlyAsp---------------570

Db   1609 TTTTAAACGTGGATTGTTGTTAAAGAGAGCTCTCCAGCAGGGTGGATCACACC 1668
              |||||||||||          ::::||||:::::
Db    571 --------------------------GlnGlnAsnProAlaProAlaGlyGlnValAspTyrThr 583

Qy   1669 ----------------------TCAGTGGGAAGAAAAATAAAATTTCCTTCAGGT 1701
                                    :::||||:::  :::
Db    584 LysAlaTrpGluGluTyrTyrLysMetGlyValAlaValProAlaProThrGly 602
```

```
Qy      854  CAGGTGTTTATCCCGCCCAGGCAGTGGGGGCCATCATCGGCAAGAAGGGGCAGCACATC  913
             |||||     :::||  :||||||||||||:||||:||     ::::|||
Db      334  GlnValLysValProArgSerThrValGlyAlaIleMetGlyLeuGlnGlySerAsnIle  353

Qy      914  AAACAGCTCTCCGGTTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAACACTGAC    973
             |||:::||||  :::|||     |||:::       ::: |||
Db      354  LysLysIleSerAsnGluThrGlyThrLysIleGlnPheMetProAspAspAspProLys  373

Qy      974  TCCAAAGTTCGTATGTTATCATCACTGGAACACGAGGCCCAATTCAAGGCTCAGGGA   1033
             |||||::     |||   :::|||||
Db      374  LeuMetGluArgThrLeuValValIleGlyAsnLysAsnLysAsnLysAsnLysValTyrValCys--Ala  392

Qy     1034  AGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAGAAGTGAAGCTG  1093
             |||:::  :::     |||     ::: ||||
Db      393  ArgLeuLeuGlnLysIleValGluAlaLysSerGluAsnAlaAsnThrProIleSerLeu  412

Qy     1094  GAGACCCACATACGTGTGCCAGCATCAGCAGCTGCCGGTCATTGGCAAAGGTGGAAAA   1153
             ::::||:||||||||:||
Db      413  --PheTyrMetLeuIleProAlaSerLysCysGlyLeuValIleIleGlyAlargGlyGlyGlu  431

Qy     1154  ACGGTGAACCAGTTGACGGCAGCTGAGGTGGTAGTACCCAAGAGAC  1207
             |||   :: :::     ::: |||
Db      432  ThrIleArgGlnIleAsnGluSerGlyAlaTyrCysGluMetSerArgAsp  449
```

RESULT 8

T02627

hypothetical protein At2g25960 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein T19L18.23
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
C;Accession: T02627; H84654
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library; August 1998
A;Description: Arabidopsis thaliana chromosome II BAC T19L18 genomic sequence.
A;Reference number: Z14681
A;Accession: T02627
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-632 <ROU>
A;Cross-references: EMBL:AC004747; NID:g3413696; PID:g3413716
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84654
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-632 <STO>
A;Cross-references: GB:AE002093; NID:g3413716; PIDN:AAC31239.1; GSPDB:GN00139
C;Genetics:
A;Gene: T19L18.23; At2g25960
A;Map position: 2
A;Introns: 112/1; 206/3; 240/3; 264/3; 304/3
C;Superfamily: glutenin

Alignment Scores:

Pred. No.:	5.03e-09	
Score:	235.50	Length: 632
Percent Similarity:	37.72%	Matches: 136
Best Local Similarity:	23.86%	Conservative: 79
Query Match:	7.57%	Mismatches: 210
DB:	2	Indels: 145
		Gaps: 24

US-09-270-437d-5 (1-1708) x T02627 (1-632)

```
Qy        4  GACGCTGCGCACCGCCGCCAGTTTACCCGGGAGGAGCCATCATGAAGTCGAATGGCCACCA   63
             ||||||     ||||||||||
Db       45  AspProThrAlaProThrLeuProSerSerTyrAsnSerValPro---ProPro        63
```

```
Qy       64  GTTGGAGAA------------------------------------CCATGCCTGAAGGTCCTACAT   96
             :::   ||                                      |||||
Db       64  MetAspGluIleGlnIleAlaLysGlnLysAlaGlnIleAlaArgLeuLeuAsn  83
```

```
Qy       97  -----------------------------------------------TGAGCAGATAGC  114
                                                              ||||||
Db       84  SerAlaAspAlaLysAlaArgProArgValAlaSerAsnGlyAlaSerTyrTyrAspTyrAspTyrAspAspAsn  103
```

```
Qy      115  ACAGGGACCTGAGAATGGGCGCCGAGGGGCCTTGGCTCTCGGGGTCAGCCCCGCCAGGG  174
             |||     :::|||||||
Db      104  LysGlyPheSerSerTyr-ProSerGluGlyLysGlnMetSerGlyThrValProSerSer  123
```

```
Qy      175  CTCACCTGGCAGCGCGGGGGGCC---CAGCCAAGCAGCAGCAAGTGACATCCCCTTCG  231
             ::::|||   |||||||  :::
Db      124  Ile-ProValSerTyrGlySerPheGlnGlyThrHisLysLysIleAspIle-------  140
```

```
Qy      232  GCTCCTGGTGCCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGAGGGGCACCATCCG  291
             |||||::||:||||||
Db      141  --------ProAsnMetArgValGlyValIleIleGlyIleGlyGlyGluThrIleLeLy  157
```

```
Qy      292  CAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGGAACGCA-----  346
             ::: :::     ||||||::: |||     :::|||     |||
Db      157  sTyrLeuGlnLeuGlnSerGlyAlaLysIleGlnValThrArgAspMetAspAlaAspPr  177
```

```
Qy      347  -GGTGCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCCGCTTG  405
             :::  :::
Db      177  oAsnCysAlaThrArgThrValAspLeuThrGlyThrProAspGlnIleSerLysAlaGl  197
```

```
Qy      406  TAAGATGATCTTGAGATTATGCATAAA---GAGGCTAAGGACACC------------------  448
             :::::||| :::   ::: ||||:::|
Db      197  uGlnLeuIleThrAspValGlnGluAlaLaGlyAsnThrAlaGlySerGlyGl  217
```

```
Qy      449  ----------------AAAACGGCTGACGAGGTTCCCCTGAAGATCCT  480
             |||:::
Db      217  yGlyGlyGlyArgMetGlyGlyGlnAlaGlyAlaAspGlnPheValMetLysIle--  236
```

```
Qy      481  GGGCCCAATAATACTTTGTGAGGCGTCTCATTGCCAAGGAAGGACGAACCTGAAGAAGGT  540
             ||||||||     ||||:::|||||:||:||||:::||||||
Db      237  ---ProAsnAsnLysValGlyLeuIleGlyLysValGlyGluThrIleLysGluSerMe  255
```

```
Qy      541  AGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTGCAA-------------GACCT  588
             :::: |||       :::     :::                    ::|||
Db      255  tGlnIleLysThrGlyAlaArgIleGlnValIleProLeuHisLeuProProGlyAspPr  275
```

```
Qy      589  TACCCTTTTACAACCCTGAGAGGACCATCACTGTGAGGGGGCCATCGAGAATTGTTGCAG  648
             ::: ||| |||     ::: :::
Db      275  oThr--------ProGluArgThrLeuGlnIleAspLyValGlyGlnIleThrGluGluHi  292
```

```
Qy      649  GGCCGAGCAGGAAATAATGAGAGAAGTTCGGGAGGGCCTATGAGAATGATGTGGCTGCCAT  708
             ||||||||     :::
Db      292  sAlaLysGlnLeuVal----------------------AsnGluIleIleSerGl  303
```

```
Qy      709  GAGCTTCCACCTGATCCCTGGCCTGAACCTGGCTGTAGGTCTTTTCCCAGCTTCATC  768
             ||| :::     |||||||::|||
Db      303  yGluAsnArgMet--------ArgAsnSerAlaMetGlyGlyTyrProGlnLnGlnGl  320
```

```
Qy      769  CAGCGGCAGTGCGCATCGGCAAGAAGGGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTAT  828
             |||::: :::     |||     :::|||||||
Db      320  yGlyTyrGlnAlaArgProProSerSer----------------Trp-AlaPro-Pro  333
```

```
Qy      829  GCAGGGCTCCGAGCAGGAGATGTGCCAGTGTTATCCCGGCCCAGGCAGT---------  879
             |||  |||
Db      334  GlyGlyProProAlaGlnProGlyTyrGlyTyrMetGlnProGlyAlaTyrProGly  353
```

```
Qy      880  --------------------GGGCGCCATCATCGGCAAGAAGGGGCAGCACAT  912
             
Db      354  ProProGlnTyrGlyTyrGlySerProGlyGlnInThrSerAlaGlyTyr  373
```

```
Qy      913  -----CAAAAGCTCTCCCGGTTTGCCAGCGCCTCCATCAAGATTGCACCACCCGAAAC  966
             |||||:::::||||
Db      374  TyrAspGlnSerSerValProProSerGlnInSer--------------------  385
```

```
Qy      967  ACCTGACTCCAAAGTTCGTATGGTTATCAT--------------------  996
```

```
Db   158 GGTCAGCCCGCCCAGGGCTCACCTGTGGCAGCGGGGCCCAGCCAGGCAAGCAGCAG------ 211
             :::||||::::         :::      ||      ::  :::
Qy    19 SerArgProGlnSerAspTyrArgAspAsnGlyGlySerLysArgGlyArgArgGlyAsp 38

Db   212 ----------------CAAGTGGACATCCCCTTGGCTCCTGGTGTGCCACC 247
                         |||
Qy    39 AspArgAspSerLeuValIleAspArgAspAspThrValPheArgTyrLeuCysProVal 58

Db   248 CAGTATGTGGTCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAG 307
             :::  ||:||||||||||         :::      |||
Qy    59 LysLysIleGlySerValIleGlyArgGlyGlyAspIleValLysGlnLeuArgAsnAsp 78

Db   308 ACCCAGTCCAAGATAGACGTGCATAGGAGGAGAACGGTGAGCTGAAAAAGCCATC 367
             ||:::||||||||         :::      |||::
Qy    79 ThrArgSerLysIleAspVal---GlyGluAlaIleProGlyCysAspArgValIle 97

Db   368 AGTGTGCACTCCACCCCT--GGTAGTACCAAGAGACCCAGACCCCTGATGAGACGACCAGGT 400
             |||                |||::    :::||:::
Qy    98 ThrIleTyrSerProAspThrThrAsnAlaPheGlyAspGluGluLysValLeuSer 117

Db   401 GCTGTAAGATGATCTTGGAGATTATGCATAAAGAG------GCTAAGGACACCAAAACG 454
             :::      ||:::
Qy   118 ProAlaGlnAspAlaLeuPheArgIleHisAspArgValAlaAspAlaArgSer 137

Db   455 GCTGAC------------------GAGGTTCCCTGAAGATCTGGCCCATAATAACTTT 496
             |||                     ||:||||||
Qy   138 GluAspSerProGluGlyLysGlnValThrAlaLysLeuLeuValProSerArgSerGln 157

Db   497 GTAGGGCGTCATTGGCAAGGAGGACGGAACCTGAAGAAGTAGAGCAAGATACCGAG 556
             :::::||||||:::         :::      |||
Qy   158 IleGlyCysIleLeuGlyArgGlyGlyAsnIleValGlnIleAsnIleArgSerGluThrGly 177

Db   557 ACAAAATACCATCCTCGTTGCCAAGACCTTACCCTT----TACAACCCTGAGAGG 610
             :::::                               |||     :::
Qy   178 AlaGlnIleArgIleArgValIleAspAspArgAsnMetProLeuCysAlaLeuAsnSerAspGlu 197

Db   611 ACCATCACTGTGAAGGGGGCCATCGAGAATTGTGCAGGGCCAGCAGAAAAATGAAG 670
             |||:::              ::  |||
Qy   198 LeuIleGlnIleSerGlyGluValLeuIleValLysLysValAlaLeuLeuGlnIleAlaSer 217

Db   671 AAAGTTCGGGAGGCCATGAGAATGATGTGGCTGCCATGAGCTCTCCACCTGATCCCTGGC 730
             :::::||||||         ::||||||
Qy   218 ArgLeu---------HisGluAsn------ProSerArgSerGlnAsnLeu--- 229

Db   731 CTGAACCTGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGCGCAGTCCCGCCGCTCCC 790
             |||||:::||||||
Qy   230 ------LeuSerSerSerGlyGlyTyrProAlaGlySerLeuMetSerHisAlaGlyGly 247

Db   791 AGCAGCGTTACTGGGCGTGCTCCTCTATAGCTCC----- 823
             :::||||||         :::
Qy   248 ProArgLeuValGlyLeuGlyLeuAlaProLeuMetGlySerTyrArgAspAlaGlyAspTrp 267

Db   824 ----------TTTATGCAGGTGCCCGCAGCAGGAG---- 856
                      :::         ::  |||::
Qy   268 SerArgProLeuTyrGlnProProArgAsnAspProAlaThrGluPheIleGluArg 287

Db   857 GTGTTTATCCCGCGCCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGGCAGCACATCAAA 916
             :::||||         :::      ||||||
Qy   288 LeuValSerProValGluAsnIleAlaSerValIleGlyLysGlyGlyAlaLeuIleAsn 307

Db   917 CAGCTCTCCCGGTTTGCCAGCGCCTTCCATCAAGATTGCACCACCCGAAACACCTGACTCC 976
             ||:||||         :::      |||::
Qy   308 GlnLeuArgGlnLeuThrArgAlaThrIleLysValAspSerArgThrGluGlyAsn 327

Db   977 AAAGTTCGTATGGTATC----------ATCACTGGACCGCCA 1009
             :::                       :::
Qy   328 AspCysLeuIleThrIleSerAlaArgGluValPheGluAspAlaTyrSerProThrIle 347

Db  1010 GAGGGCCAATTCAAGGCTCAGGGAAGAAATCATGGCAAACTCAAGGAGGAGAACTTCTTT 1069
             ||||||         :::
Qy   348 GluAlaValMetArgLeuGlnProLysCysSerAspLysValGlnArgAspSerGlyLeu 367

Db  1070 GGTCCCAAGGAGGAGAAGCTGAAGCTGCAGACCCACATACGTGTGCCAGCATCAGCAGCTGGC 1129
```

```
Db   386 ------AlaGlnGlyGluPheGlyTyrTyrGlyValGlnGlnSerGlnProSerSer 403
             |||                      |||      :::
Qy   997 CACTGGACGCACGAGGCCCAATTCAAGGCTCAGGGAAGAAATCATGGCAAACTCAAGGA 1056

Db   404 GlyGlySerSerAlaProProThrAspThrTyrThrGlyTyrAsnTyrTyrGlnHisAlaSer 423
             :::::      |||::
Qy  1057 GGAGAACTTCTTTGGTCCCAAGGAGGAGAAGCTGAAGCTGCAGACCCACACGTGTGCCAGC 1116

Db   424 GlyTyrGlyGlnAlaGlyGlyTyrGlnGln------AspGlyTyrGlyAlaTyrAsn 441
             ||||||                :::
Qy  1117 ATCAGCAGCTGGCCCGGGTCATTGGCAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTT 1176

Db   442 AlaSerGlnGlnSerGlyTyrGlyGln------AlaAlaGlyTyr 454
             |||::         :::
Qy  1177 GACGGCAGCAGTGAGGT--GGTAGTACCAAGAGACCCAGACCCCTGATGAGACGACCAGGT 1233

Db   455 AspGlnGlnGlyGlyTyrGlySerThrThrAsnPro-SerGlnGluGluAsp------- 471
             ||:::      |||
Qy  1234 CATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGA 1293

Db   472 ------AlaSerGlnAlaAlaPro------------------ 477
             |||
Qy  1294 CATCCTGGCCCAGGTTAAGCAGCGATCAGAAGGGACAGAGTAACCAGGCCCCAGGCACG 1353

Db   478 ------------ProSerAlaGlnSerGlyGlnAlaGlyTyrGlyThrThrGly 492
             |||||:::
Qy  1354 GAGGAAGTGACCAGCCCCTCCCTGTCCCTTNGAGTCCAGGACAACAACGGGCGAGAAATCG 1413

Db   492 ValGlnProProAlaGlnGlySerThrGlyValGlnAlaGlyTyrGlyAlaProProThrSe 512
             ||||||::::::
Qy  1414 AGAGTTGCTCTCCCCGGCAGGCCTGAGAATGAGTGGAATCCGGGACACNTGGCGCCGGG 1473

Db   512 rGlnAlaGlyTyrSerSerGlnProAlaAlaAlaAlaTyrAsnSerGlyTyrGlyAlaProPr 532
             :::||||||
Qy  1474 CTGTAGATCAGGTTTGCCCACT 1495

Db   532 oProAlaSerLysProProThr 539
             ||||||
Qy  1474 CTGTAGATCAGGTTTGCCCACT 1495
```

RESULT 9

T49962

hypothetical protein F8M21.160 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49962
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24493
A;Accession: T49962
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-568 <BEV>
A;Cross-references: EMBL:AL353993; GSPDB:GN00063; ATSP:F8M21.160
A;Experimental source: cultivar Columbia; BAC clone F8M21
C;Genetics:
A;Gene: ATSP:F8M21.160
A;Map position: 5
A;Introns: 200/3; 337/3; 544/3

Alignment Scores:
Pred. No.:   568
Score: 5.78e-09
Percent Similarity: 234.50   Length: 568
Best Local Similarity: 42.47%   Matches: 92
Query Match: 20.67%   Conservative: 97
DB: 7.54%   Mismatches: 175
                           2   Indels: 81
                                    Gaps: 16

US-09-270-437D-5 (1-1708) x T49962 (1-568)

```
Qy    98 CCCGATGAGCGAGATAGCACAGGACCTGAGATGTGGCGGCGCGGGGGGGCTTTGGCTCTCGG 157
             ||||||         :::
Db     3 ProAspHisArgMetSerProAsphisArgAspSerHisArg----------LysArg 18

Qy   158 GGTCCCAAGGAGGAGAAGCTGAAGCTGCAGACCCCACATACGTGTGCCAGCATCAGCAGCTGGC 1129
```

```
Db   569  ATTCCTCGTTGCAAGAC---CTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAG  625
          :::          ||||||  ::::::: ||||||||||
Qy   168  Leu------ValGlnAspAsnGlnSerValSerLysProLeuArgIleThr  185

Qy   626  GGGGCCATCGAGAATTGTTGCAGGGCCAGCAGTTGCAGAATTTGCAGCTGAG  1189
          :::||||||||||:|||||||||||          ||:|:::   :: ::::
Db   186  GlyAspProGlnLysIleLysGlyAlaIleIleThrGluMetArgArgMetThrLysAlaAsn  205

Qy   686  TATGAGAATGATGTGGCTGCCATGAGCTCCACTGAGCTCCTGGCCTGGCTGCT  745
          :::   :::::: ::::
Db   206  GlyAspGlyAsnGlyGlySer------------GlyLeuGlnMetHisHis  218

Qy   746  GTAGGTCTTTTCCCAGCTTCATCCAGCGAGGCCTCCGGCCCTCCCAGCCAGCGTTACTGGG  805
          |||                                              |||
Db   219  AlaGly------------------------------------------Gly  221

Qy   806  GCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGGAGCAGGAGGTGTGCAGGTGTTTATC  865
          ||:::|                                    ::||||
Db   222  GlyGlyGlyAlaSerAla------------------------ArgGlyGluValValVal  233

Qy   866  CCCGCCAGGCAGTGGGGCGCCATCATGGCAGGAGAGGGGCAGCACATCAAACAGCCTCC  925
          ||||||||||||||||||||||||:|||||||||||:::::||||||:::
Db   234  ProArgSerValGlyIleIleIleIleIleGlyGlnGlyAspThrIleLysArgLeuAla  253

Qy   926  CGGTTTGCCAGCGCCCTCCATCAAGATGCACCACCCGAAACACCTGACTCCAAAGTTCGT  985
          |||||||                            :: :::         ||||||
Db   254  MetGluThrGlyThrLysIleGlnPheLysProAspAspProSerThrProGluArg  273

Qy   986  ATGGTTATCATCACTGGACCGCCCAGAGGCCCAATTCAGGCTCAGGGAGAATCATGGC  1045
          ::||| |||||      :: :||||||    :::::||||||||
Db   274  CysAlaValIleMetGlyThrArgAspArgIleTyrArgAlaThrGluArgIleThrGlu  293

Qy   1046 AAACTCAAGGAGGAG---------------------AACTTCTTGGTCCCAGGAG  1081
          ::||||||:|:|:||
Db   294  LeuValIleLysSerThrMetGlnGlnGlyGlyGlyGlyAsnValAlaGlyAlaMetVal  313

Qy   1082 GAAGTGAAGCTGAGACC---CACATACGTGTGCCAGCATCAGCAGCTGGCCGGGTCATT  1138
          |||  ::::: |||||||:::                ||||||| |||||||||||
Db   314  SerAsnGluAlaSerThrPheTyrMetSerValProAlaAlaLysGlyLeuValIle  333

Qy   1139 GGCAAAGGTGAAAAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTA  1198
          ||:||||||||||||||| ::: |||    :: :::|||||||
Db   334  GlyLysGlyLysGlyGluThrValLysLysIleAsnSerGlueSerGlyAlaHisCysGluLeu  353

Qy   1199 CCAAGAGACCAGCAGCCCTGATGAGACGACCAGGTCATCGTG----------AAA  1243
          |||||||:::                        :::::::::
Db   354  SerArgAspProThrGlyAsnAlaAspGluLysValPheValIleLysGlyGlyLysLysArg  373

Qy   1244 ATCATCGGACATTCTATGCCAGTCAGATGGCTCAACGGAGACATCCTGGCC  1303
          |||                :::   |||  :::  ||||||
Db   374  AlaIleGluHis------AlaLysHisLeuPheLeuArgIleLysValAlaAspIleAlaPro  391

Qy   1304 CAGGTTAAGCACGCAGCATCAAGAGGGACCAGAGTAACCAGGCCACCGGAGGAGTGA  1363
          ::: :::      :::      :::
Db   392  AsnThrProPheArgAspAspSerAlaMetThrMetGlnThrGlnPhe-------  407

Qy   1364 CCAGCCCCTCCTGTCCCTTNGAGTCCAGGCAGCAACACGGGCAGAAATCAGAGAGTGTGCT  1423
          :::                     ||||||
Db   408  ---SerAlaProAlaGlnAlaAsnAsnPheGlyGlyGlnGlnGlnTrpAsnProValAlaGln  426

Qy   1424 CTCCCCGGCAGGCCTGAGAATGAGTGGAATCCGGG-----------  1459
          ::|||||                     ||
Db   427  IleProAlaAlaAlaGlnAsnProTyrGlnValGlyLysTrpGlnAsnSerValTyr  446

Qy   1460 ------------------------ACACNTGGGCCCGGCTGTAGATCACGTTTG----  1489

Qy   1490 ------------CCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACC  1525
                      ::|||||
Db   467  GlnGlnGlnValAlaAlaTyrGlnGlnProGlnValGlnGlnVal---GlnProAlaThr  485

Qy   1526 CTGATCTNTCAGCCCCCAAACACCCACC----------  1552
```

Left block:

```
Db   368  -----------ValSerPheThrThrArgLeuLeuValProSerSerArgIleGly  382

Qy   1130 CGGGTCATTGGCAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTTGCAGCTGAGCAGCTGAG  1189
          :::||||||||||||||||||||          ||:|:::
Db   383  CysIleLeuGlyLysGlyGlyAlaIleIleThrGluMetArgArgMetThrLysAlaAsn  402

Qy   1190 GTG--GTAGTACCAAGAGACCAGACCCT------------GATGAGAACGACCAGGTC  1234
          :::  ::::::  :::::::::::        ||||||
Db   403  IleArgIleLeuGlyLysGluAsnLeuProLysValAlaSerAspAspAspGluMetVal  422

Qy   1235 ATCGTGAAAATCATC  1249
          ::: :::
Db   423  GlnValAsnPheMet  427
```

RESULT 10

T19216

hypothetical protein C12D8.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19216
R;McMurray, A.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19092
A;Accession: T19216
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-589 <WIL>
A;Cross-references: EMBL:Z73969; PIDN:CAA98232.1; GSPDB:GN00023; CESP:C12D8.1a
A;Experimental source: clone C12D8
C;Genetics:
A;Gene: CESP:C12D8.1a
A;Map position: 5
A;Introns: 7/3; 25/1; 159/3; 318/3; 513/2

Alignment Scores:

Pred. No.:	8.86e-09
Score:	232.00
Percent Similarity:	37.15%
Best Local Similarity:	18.75%
Query Match:	7.46%
DB:	2

Length:	589
Matches:	108
Conservative:	106
Mismatches:	232
Indels:	130
Gaps:	18

US-09-270-437D-5 (1-1708) x T19216 (1-589)

```
Qy   197  CCAGGCAAGCAGCAGCAAGTGGACATCCCCTT----------------------  229
          ||||||||:::                ||||
Db   28   ProAlaLysArgProAlaAspAspThrAspLeuAsnProPheMetAspAspAsnGluAla  47

Qy   230  --------CGGCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGG  280
          ::        |||
Db   48   ValAsnGluLysTyrProIleProGluserAlaValGlyIleIleValGlyIleArgGlyGly  67

Qy   281  GCCACCATCGCAACATCACAAAAACAGACCCCAGTCCAGATAGCACCGTCCATAGGAAGGAG  340
          ::: ||||:::        :::     :::
Db   68   SerGluIleGlnGlyIleGlnAlaLysGlyAlaGlyCysValGlnMetSerProAspAla  87

Qy   341  AACGCAGGTTGCAGTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCC  400
          |||||||:::                       ::|||||
Db   88   AspProSerGlyValArgMetValThrLeuGluLeuGlySerArgSerArgSerAsnValGluThr  107

Qy   401  GCTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAA---------  451
          |||||:::        :::  :::
Db   108  AlaLysHisLeuIleLeuAsnGluValValAlaLysArgSerGlnAsnProArgGlnInTyrGly  127

Qy   452  --ACGGCTCGACGAGGTTCCCTGAAGATCCTGGCCCATAATAACTTGTAGGGCGTCTC  508
          ::  ||||       :::     ::: ::
Db   128  PheProArgAlaGlnThrThrIleAspIleProProAsnArgCysGlyLeuIle  147

Qy   509  ATTGGCAAGGAAGGACGGAAGAACCTGAAGAAGTTAGAGCAAGATACCGAGACAAAATCACC  568
          ::: |||||||||||||    ::||||||       :::
Db   148  IleGlyLysSerGlyAspThrLysArgGlnInLeuGlnLysGlnLysLysSerCysLysMetIle  167
```

```
Db   486 ThrValAlaAlaThrThrProThrValAspProValThrGlyGluGlnAspTyrSer 505
Qy  1553 --CAATTGGCCACACTGTNTGCCCCTCGGGGTG------------------------ 1585
Db   506 AlaGlnTrpGluTyrTyrLysSerIleGlyAlaHisAspLysAlaGluAlaValGlu 525
Qy  1586 --------------------TCAGAAATNTAGCGCAGGCACTTTAAACGTGGATTG 1624
Db   526 AlaGlnMetLysLysLysAlaGluAlaAlaArgAlaValProGlyGlyLeu 545
Qy  1625 TTTAAAGAAGCTCTCCAGGCCCCCACCAAGAGGGTCGGATCACACCTCAG 1672
Db   546 --------------IleGlnGlnMetProMetGlyMetAlaMetProGln 557
```

RESULT 11
T19217
hypothetical protein C12D8.1b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19217
R;McMurray, A.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19092
A;Accession: T19217
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-611 <WIL>
A;Cross-references: EMBL:Z73969; PIDN:CAA98233.1; GSPDB:GN00023; CESP:C12D8.1b
A;Experimental source: clone C12D8
C;Genetics:
A;Gene: CESP:C12D8.1b
A;Map position: 5
A;Introns: 181/3; 340/3; 535/2

Alignment Scores:
Pred. No.:   8.95e-09
Score: 232.00
Percent Similarity: 37.15%
Best Local Similarity: 18.75%
Query Match: 7.46%
DB:

Length: 611
Matches: 108
Conservative: 106
Mismatches: 232
Indels: 130
Gaps: 18

US-09-270-437D-5 (1-1708) x T19217 (1-611)

```
Qy   197 CCAGCCAAGCAGCAGAAGTGGACATCCCCTT------------------------ 229
Db    50 ProAlaLysArgProAlaAspThrAspLeuAsnProPheMetAspAspAsnGluAla 69
Qy   230 --------CGGCTCCTGGTGCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGG 280
Db    70 ValAsnGluLysTyrProIleProGluSerAlaValGlyIleValIleGlyArgGlyGly 89
Qy   281 GCCACCATCCGCAACATCACAAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAG 340
Db    90 SerGluIleGlnGlyIleGlnAlaLysAlaGlyCysArgValGlnMetSerProAspAla 109
Qy   341 AACGGCAGGTGCAGTGGAAAAAGCCATCAGTGCACTCCACCCCTGAGGGCTGCTCCTCC 400
Db   110 AspProSerSerGlyValArgMetValThrLeuGluGlySerArgSerAsnValGluThr 129
Qy   401 GCTTTGTAAGATGATCTTGGAGATTATCCATAAAGAGGCTAAGGACACCAAA-------- 451
Db   130 AlaLysHisLeuIleAsnGluValAlaValAlaArgSerGlnAsnProArgProGlnTyrGly 149
Qy   452 ---ACCGGCTGACGAGGAGGACGAACCTGAAGAGGTAGAGCCAAGATACCGAGACAAAATCACC 508
Db   150 PheProArgAlaGlnThrThrIleAspIleAlaIleProProAsnArgCysArgGlyLeuIle 169
Qy   509 ATTGGCAAGGAAGGACGCCGAACCTGAGAAGGTAGAGCGCAAGGAGCACGAGAGAACC 568
Db   170 IleGlyLysSerGlyAspThrIleArgGlnLeuGlnLeuGlnArgGlyLysCysGlyLysMetIle 189
```

```
Qy   569 ATCTCCTCGTTGCAAGAC---CTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAG 625
Db   190 Leu------ValGlnAspAsnGlnSerValSerAspGlnSerLysProLeuArgIleThr 207
Qy   626 GGGGCCATCGAAGAATTGTTGCAGGGCCAGCAGGAAATGAAGAAAGTTCGGGAGGCC 685
Db   208 GlyAspProGlnLysGluLeuAlaHisLeuAlaLysGlnLeuLeuValAlaGluIleLeuAsnSerGly 227
Qy   686 TATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGATCCTGCCTGAACCTGGCTGCT 745
Db   228 GlyAspGlyAsnGlyGlySer-----------------------GlyLeuGlnMetHisHis 240
Qy   746 GTAGGTCTTTTCCCAGCTTCATCCAGCGAGTCCGCGCCTCCCAGCAGCGTTACTGGG 805
Db   241 AlaGly---------------------------------------------------Gly 243
Qy   806 GCTGCTCCCTATAGTCCTTCCTTTATGCAGGCTCCCAGCAGCAGGAGATGGTGCAGGTGTTTATC 865
Db   244 GlyGlyGlyAlaSerAla----------------------------ArgGlyGluValVal 255
Qy   866 CCCGCCCAGCAGTGGGCGCCATCATCGGCCAAGAGAGGGCAGCACATCAAACAGCTCTCC 925
Db   256 ProArgSerSerValGlyIleIleIleGlyLysGlnGlyGlyAspThrIleLysArgLeuAla 275
Qy   926 CGGTTTGCCAGCGCCTCCATCAAGATTGCACCACCCAGACAACCACTGACTCCAAAGTTCGT 985
Db   276 MetGluThrGlyThrLysIleGlnIlePheLysProAspAspAspProSerThrProGluArg 295
Qy   986 ATGGTTATCATCACTGGACCGCCGCCAGAGGCCCAATTCAAGGCTCAGGGAGGAGAATCATGGC 1045
Db   296 CysAlaValIleMetGlyThrArgAspGlnIleTyrArgAlaThrGluArgIleGluThrGlu 315
Qy  1046 AAACTCAAGGAGGAG-------------------------AACTTCTTTGGTCCCAAGGAG 1081
Db   316 LeuValLysGlySerThrMetGlnIngIleGlyIleGlyIleGlyIleGlyAsnValAlaGlyAlaMetVal 335
Qy  1082 GAAGTGAAGCTGGAGACC---CACATACGTGTGCCAGCATCAGCAGCTGCGGGTCATT 1138
Db   336 SerAsnGluAlaSerThrPheTyrMetSerValProAlaAlaLysCysGlyLeuValIle 355
Qy  1139 GGCAAGGTGGAAAAACGGTGAACAGTTGCAGAATTTGACGGCAGCTGAGGTGTAGTA 1198
Db   356 GlyIleGlyIleGlyThrIleLysGlnIleLysAsnSerGlnAlaHisCysGlyLeu 375
Qy  1199 CCAAGAGACCAGACCCCTGATGAGAACGACCAGTCATCGTG----------------AAA 1243
Db   376 SerArgArgAspProThrGlyAsnAlaAspGluLysValPheValIleLysGlyLysLysArg 395
Qy  1244 ATCATCCGACATTCTATGCCAGTCGAGATGGCTCAACGGAAGATCCGAGACATCCTGGCC 1303
Db   396 AlaIleGluHis------AlaLysHisLeuIleArgIleLysValGlyAspIleAlaPro 413
Qy  1304 CAGGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACCAGGCCACGGAGGAGTGA----------- 1363
Db   414 AsnThrProPheArgAspAspSerAlaMetThrMetGlnThrGlnPhe---------- 429
Qy  1364 CCAGCCCCTCCGTGTCCCTTNGAGTCCAGGACAACAACGGGCAGAAATCGAGAGTGTGCT 1423
Db   430 ---SerAlaProAlaGlnAsnAsnPheGlyGlyGlnGlnTrpAsnProValAlaGln 448
Qy  1424 CTCCCGGCCAGGCCTGAGAATGAGTGGGAATCCGGG------------------------ 1459
Db   449 IleProAlaAlaAlaGlnAsnProTyrGlnValGlyGlyTrpGlnGlnAsnSerValTyr 468
Qy  1460 ---------------------------------ACACNTGGCCGGCTGTAGATCAGGTTTG- 1489
Db   469 AlaGlnThrAlaAlaProAlaAlaAlaProTyrAlaAlaAlaGlyIleValGlnInPro 488
Qy  1490 ---------------------------CCCACTTGATTGAGAAGATGTTCCAGTGAGGAACC 1525
Db   489 GlnIngGlnValAlaTyrGlnGlnInProGlnValGlnGlnInVal---GlnProAlaThr 507
Qy  1526 CTGATCTNTCAGCCCCAAACACCCACC------------------------------- 1552
```

```
Db    508  ThrValAlaAlaThrThrProThrValAspProValThrGlyGluGlnAspTyrSer  527
                :::  |||||||
Qy   1553  ---CAATTGGCCCAACACTGNTGCCCCTCGGGGTG------------------------  1585
                :::|||
Db    528  AlaGlnTrpMetGluTyrTyrLysSerIleGlyTyrLysSerIleGlyAlaHisAspLysAlaGluAlaValGlu  547
Qy   1586  --------------TCAGAAATTNTAGCCAAGGCACTTTTAAACGTGGATTG  1624
                                    ::|||||||  |||
Db    548  AlaGlnMetLysLysLysLysAlaGluAlaAlaAlaAlaArgAlaValAlaValProGlyGlyLeu  567
Qy   1625  TTTAAAGAGAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAG  1672
                   ::|||    |||  |||  ||||||
Db    568  ----------IleGlnGlnMetProMetGlyMetAlaMetProGln  579

RESULT 12
I38489
onconeural ventral antigen-1 - human
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999
C;Accession: I38489
R;Buckanovich, R.J.; Posner, J.B.; Darnell, R.B.
Neuron 11, 657-672, 1993
A;Title: Nova, the Paraneoplastic Ri Antigen, is Homologous to an RNA-Binding Protein an
A;Reference number: I38489; MUID:94000830; PMID:8398153
A;Accession: I38489
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-510 <RES>
A;Cross-references: EMBL:U04840; NID:g440877; PIDN:AAA16022.1; PID:g440878
C;Genetics:
A;Gene: Nova-1

Alignment Scores:
Pred. No.:            4.92e-08
Score:               221.50
Percent Similarity:   34.89%          Length:       510
Best Local Similarity: 21.64%         Matches:      111
Query Match:          7.12%           Conservative:  68
DB:                   2               Mismatches:   173
                                      Indels:       161
                                      Gaps:          17

US-09-270-437D-5 (1-1708) x I38489 (1-510)

Qy    125  GAGAATGGGCGCGCGAGGGGGCTTT------------------------GGCTCTCGGGGTCAG  163
                ::|||||||||                                :: |||
Db      9  GlnAsnGlyThrHisThrGlyValProIleAspLeuAspProProAspSerArgLysArg  28

Qy    164  CCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCAAGGCCCAAGCCAAGTGGAC---  220
                |||   :: ||||||||                           ||||:
Db     29  ProLeuGluAlaProGlyGluAlaProGlyThrSerArgThrAsnThrGlyGluAspGly  48

Qy    221  ---ATCCCCTTCGGCTCCTGTGCCCACCCAGTATGTGGTGCCATTATTGGCAAGAGAG  277
                   |||:::::|||||||                      |||||||||||
Db     49  GlnTyrPheLeuLysValLeuValLeuIleProSerTyrAlaAlaGlySerIleLeuIleGlyLysGly  68

Qy    278  GGGGCCACCATCCGCAACATCCAACACAAAACACCAGTCCAAGATAGACGTGCATAGGAAG  337
                   :::|||                                        :: :::
Db     69  GlyGlnThrIleValGlnIleAsnGlyHisGlyAlaThrGlyAlaThrIleLysLysSerLysLysLeu  88

Qy    338  GAGAACGCA------------GGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACC  382
                :::  |||                   |||:||
Db     89  SerLysSerLysAspPheTyrProGlyThrGlyThrArgValCysLeuIleGlnGlyThr  108

Qy    383  CCTGAGGGGCTGCTCCTCCCGCT-----------TGTAAGATGATCTTGGAGATTATCAT  430
                |||:::::|||:||:||                ||:  |||
Db    109  ValGluAlaLeuAsnAlaValHisGlyPheIleAlaGlyLysIleArgGluMetProGln  128

Qy    431  AAAGAGGTAAG------------------------GACACCAAACGGCTGACGAGGTTCCC  469
                |||:|:::::                      ||||||:::::::
Db    129  AsnValAlaLysThrGluProValSerIleLeuGlnProGlnThrValAsnProAsp  148
```

```
Db    149  ArgIleLysGlnThrLeuProSerProThrThrLysSerProSerProSerAspPro  168
                                                     |||||||||
Qy    470  CTG---------------AAGATCCTGGCCACCAATAATAACTTTGTAGGGGCT  505
                :::
Db    169  MetThrThrSerArgAlaAsnGlnInValLysIleIleValProAsnSerThrAlaGlyTyrLeu  188
Qy    506  CTCATTGGCCAAGGAAGGACGGAACCTGAGAAGGTGAGACAAGATACCCAGACAAAATC  565
                :::|||||||  |||   |||  :: |||
Db    189  IleIleGlyLysLysGlyGlyAlaThrValLysAlaValMetGluGlnSerGlyAlaTrpVal  208
Qy    566  ACCATCTCCTCGTTGCAAGACCTTACCCTTACAACCTGAGAGGACCATCACTGTGAAG  625
                |||:::||||
Db    209  GlnLeuSerGlnLysProAsp---GlyIleAsnLeuGlnLysGlnArgValValThrValSer  227
Qy    626  GGGGCCATCGAGAATTGTTGCACGGCCCAGACGCAGCAAATAATGAGAAAGTTCGGGAG---  682
                |||           :::   |||||||
Db    228  GlyGluProGluGlnAsnArgLysAlaValGluLeuIleIleGluGlnLysIleGluGluAsp  247
Qy    683  ------------------------------GCCATGAGAAT----  694
                                              :::|||
Db    248  ProGlnSerGlySerCysLeuAsnIleSerTyrAlaAsnValThrGlyProValAlaAsn  267
Qy    695  --------------------------GATGTGGCTGCCATG  709
                                          ::: |||
Db    268  SerAsnProThrGlySerProTyrAlaAsnThrAlaGlyValLeuProThrAlaAlaAla  287
Qy    710  AGCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGAGGTCTTTCCCA----------  760
                :::::|||   |||:::                   ||||||
Db    288  AlaAlaGlyLeuLeuGlyHisAlaAsnLeuAlaLeuAlaGlyValAlaAlaAlaPheProAlaValLeu  307
Qy    760  -----------------------------------------------------------  760
Db    308  SerGlyPheThrGlyAsnAspLeuValAlaIleAlaThrSerAlaLeuAsnThrLeuAlaSer  327
Qy    761  ----------------------------------------GCTTCATCCAGCGCA  775
                                                        |||||:||
Db    328  TyrGlyTyrAsnThrLeuAsnThrLeuGluGlyLeuGlyLeuSerGlnAlaAlaAlaThrArgGlyAla  347
Qy    776  GTCCCGCCGCCTCCCAGCAGGTTACTGGGGCTGCTCCCATAGCTCCTTTATGCAGGCT  835
                :::|||                GATGTGGCTGCCATG
Db    348  LeuAlaAlaAlaAlaAlaSerAlaAsnProAlaAlaAlaAlaAlaAsnLeuLeuAlaThr  367
Qy    836  CCCGAGGAGGAGGATGGTGCAGGTGTTTATCCCCGGCCAGTGGGCGGCCATCATCCGGC  895
                                                          AlaSerAlaSerGlySerThrAlaGly
Db    368  TyrAlaSerGlu----------------------------------  380
Qy    896  AAGAAGGGGCAGCACATCAAACAGCTCTCCGGTTGCCAGCGCCTCCATCAAGATTGCA  955
                |||||:||                           |||:::
Db    381  GlyThrAla----------GlyThrPheAlaLeuGlySerLeuAlaAlaAla  394
Qy    956  CCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGACGCGCCAGAGGCC  1015
                                          ThrAlaAlaThrAsnGly
Db    395  --------------------------  400
Qy   1016  CAATTCAAGGCTCAGGAGAAGAATCATGGCAAACTC-----AAGGAGGAGAACTTC  1066
                                                     |||||:
Db    401  TyrPheGlyAlaAlaSerProLeuAlaAlaSerAlaAlaIleLeuGlyThrGluLysSerThr  420
Qy   1067  TTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACCGTGCCAGCATCAGCAGCT  1126
                ||||:: |||||:                       IleAlaValProGluAsnLeuVal
Db    421  AspGlySerLysLysAspValGlu----------  436
Qy   1127  GGGCCGGTCATTGGACAAAGGTGGAAAAAACGGTGAACGAGTTGACGGCAGCT  1186
                |||  :::::||||||||||                ||||
Db    437  GlyAlaIleLeuGlyLysGlyGlyLysThrLeuValGluLeuGlnGlyLeuThrGlyAla  456
Qy   1187  GAGGTGGTAGTACCAAGA------GACCAGACCCCTGATGAGAACGACCAGGTCATCGTG  1240
                |||  :::                 :::
Db    457  ArgIleGlnIleSerLysGlyGlyGluPheValProGlyThrArgAsnProAspLys---Val  475
Qy   1241  AAAATCATCCGACATTTCATGCCAGTCAGTGACATGGCTCAA  1279
                ||||||||
Db    476  ThrIleThrGlyThrProAlaAlaThrGlnAlaAlaGlnAlaAlaGln  488
```

RESULT 13
S58529
alpha-complex protein 1 - human
N;Alternate names: nucleic acid-binding protein; protein PCBP-1
C;Species: Homo sapiens (man)
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 01-Dec-2000
C;Accession: S58529; S58523; S65678; S43489; S41378; S42472
R;Kiledjian, M.; Wang, X.; Liebhaber, S.A.
EMBO J. 14, 4357-4364, 1995
A;Title: Identification of two KH domain proteins in the alpha-globin mRNP stability com
A;Reference number: S58523; MUID:96016208; PMID:7556077
A;Accession: S58529
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-356 <KIL>
A;Accession: S58523
A;Status: preliminary
A;Molecule type: protein
A;Residues: 125-139;251-265;315-323 <KI2>
R;Leffers, H.; Dejgaard, K.; Celis, J.E.
Eur. J. Biochem. 230, 447-453, 1995
A;Title: Characterisation of two major cellular poly(rC)-binding human proteins, each co
A;Reference number: S65678; MUID:95331278; PMID:7607214
A;Accession: S65678
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-204,'V',206-356 <LEF>
A;Cross-references: EMBL.X78137; NID:g460770; PIDN:CAA55016.1; PID:g460771
A;Experimental source: AMA cells (transformed human amnion cells)
A;Note: submitted to the EMBL Data Library, March 1994
R;Aasheim, H.C.; Loukianova, T.; Deggerdal, A.; Smeland, E.B.
Nucleic Acids Res. 22, 959-964, 1994
A;Title: Tissue specific expression and cDNA structure of a human transcript encoding a
A;Reference number: S43489; MUID:94203810; PMID:8152927
A;Accession: S43489
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-298,'H' <AAS>
A;Cross-references: EMBL:Z29505; NID:g444020; PIDN:CAA82631.1; PID:g444021
C;Genetics:
A;Gene: GDB:HNRNPX
A;Cross-references: GDB:344947
C;Keywords: RNA binding

Alignment Scores:
Pred. No.:   7.36e-08
Score: 218.50
Percent Similarity: 41.90%
Best Local Similarity: 20.95%
Query Match: 7.03%
DB: 2

Length: 356
Matches: 75
Conservative: 75
Mismatches: 139
Indels: 69
Gaps: 9

US-09-270-437D-5 (1-1708) x S58529 (1-356)

```
Qy    188 GGCGGGGGCCCAGCCAGCAGGCAGCAGCAGTGGACAATCCCCTTCGGCTCCTGGTGCCACC 247
              ||||||                  ||          ::||||||:||:|| ::||||
Db      3 AlaGlyValThrGluSerGlyLeuAsnValThrLeuThrIleArgLeuLeuMetHisGly 22

Qy    248 CAGTATGTGGGTGCCATTATTGCCAAGGAGGGGCCACCATCCGCAACATCACAAAACAG 307
             ::: ||||||::|||||::|||              ::::::::::   ::::::
Db     23 LysGluValGlySerIleIleGlyLysGlyGlySerValLysArgIleArgGluGlu 42

Qy    308 ACCCAGTCCAAGATAGACGTGCATAGGAAGGAGGAAGACCCAGGTGCAGCTGAAAAGGCCATC 367
             ||||::|:::::                                 ||||||||||::
Db     43 SerGlyAlaArgIleAsnIle---------SerGluGlyAsnCysProGluArgIleIle 59

Qy    368 AGTGTGCACTCACCCCTGAGGCGGCTGCTCCTCCGCTTGTGTAAGATGANTCTTGAGAGATTATG 427
                                            :::::|| ::::: ||
Db     60 ThrLeuThrGlyProThrAsnAlaIlePheLysAlaPheAlaMetIleIleAspLysLeu 79

Qy    428 CATAAAGAGGCTAAGGACACCAAAACG---------------GCTGACGAGGTTCCC 469
             ::::::           :::|| |||
```

RESULT 14
S41224
hnRNP protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C;Accession: S41224
R;Siomi, H.; Matunis, M.J.; Michael, W.M.; Dreyfuss, G.
Nucleic Acids Res. 21, 1193-1198, 1993
A;Title: The pre-mRNA binding K protein contains a novel evolutionarily conserved motif
A;Reference number: S41224; MUID:93219080; PMID:8464704
A;Accession: S41224
A;Molecule type: mRNA
A;Residues: 1-396 <SIO>
C;Superfamily: transformation upregulated nuclear protein HNRPK

Alignment Scores:
Pred. No.:   8.95e-08
Score: 217.50
Percent Similarity: 41.67%

Length: 396
Matches: 90
Conservative: 75

```
Db     80 GluGluAspIleAsnSerSerMetThrAsnSerThrAlaAlaSerArgProProValThr 99

Qy    470 CTGAAGATCCTGCCCATATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAAC 529
              ||||||:::::::::::                ||||||:::|||||||||
Db    100 LeuArgLeuValProAlaThrGlnCysGlySerLeuIleGlySerLeuGlyCysLys 119

Qy    530 CTGGAAGAGTAGACAAGATACCGGACAAGAAAATCACCATCTCCGTTGCAAGACCTT 589
             ::||:|||:::: ||           :::::|||           ------AspMet 137
Db    120 IleIleLysGluIleArgGluSerThrGlyAlaGlnValAlaGly---

Qy    590 ACCCTTACAACCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGCAGG 649
             ||||||:| :::::::::         ||||||||
Db    138 LeuProAsnSerThrGluArgGlyAlaIleThrIleAlaGlyValProGlnSerValThrGlu 157

Qy    650 GCCGAGCAGGAAATA---------ATGAAGAAGTTCGGAGGCCATGAGAAGAATGAT 697
             ||||||||||:::            :::::::::||
Db    158 CysValLysGlnIleCysLeuValMetGluThrLeuSerGlnSerProGlnGlyArg 177

Qy    698 GTGGCTGCCATGAGCTCTCACCTGATCCCTGGCCTGAAC-----CTGGCTGCTGTAGGT 751
             ||:::                   :::||           :::
Db    178 ValMetThrIleProTyrGlnProMetProAlaSerSerProValIleCysGlyAlaGlyGly 197

Qy    752 CTT---------------TTCCAGCTTCATCCAGCGACAGTCCCGCCGCCT 787
                               :::::::::::::::
Db    198 GlnAspArgCysSerAspAlaAlaGlyTyrProHisAlaThrHisAspLeuGluGlyPro 217

Qy    788 CCC-------------------------------------------------- 790
             |||
Db    218 ProLeuAspAlaTyrSerIleGlnGlyGlnHisThrIleSerProLeuAspLeuAlaLys 237

Qy    790 ---------------------------------------------------- 790

Db    238 LeuAsnGlnValAlaArgGlnGlnSerHisPheAlaMetMetHisGlyThrGlyPhe 257

Qy    791 AGCAGCGTTACTGGGGCTGCTCCC---------------TATAGTCCTTTATGCAGGCTCCC 838
             :::     ||||||::|                   ||||||||||
Db    258 AlaGlyIleAspSerSerProGluValLysGlyTyrTrpAlaSerLeuAspAlaSer 277

Qy    839 GAGCAGGAGATGGTGCAGGTGTTTATCCCGCCCAGGCAGTGGGGCCATCATCGGCAAG 898
             |||          :::::: |||||           ::||
Db    278 ThrGlnThrThrHisGluLeuThrIleProAsnAsnLeuIleGlyCysIleLeuGlyArg 297

Qy    899 AAGGGGCCAGCACATCAAAACAGCTCTCCCGGTTTGCCAGCGGCCTCCATCAAGATTGCACCA 958
             :::||        :::::::|| :::::::                   ::
Db    298 GlnGlyAlaAsnIleAsnGluIleArgGlnMetSerGlyAlaGlnIleLysIleAlaAsn 317

Qy    959 CCCGAAACACCTGACTCCAAAGTTCGTATCGTATCGTTATCATCACTGGACCGCCGAGAGCCCAA 1018
             |||                   ::: :::::::
Db    318 Pro---ValGluGlySerGlyArgGlnValThrIleThrGlySerAlaAlaSerIle 336

Qy   1019 TTCAAGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGT 1072
             |||||                         ::|||||
Db    337 SerLeuAlaGlnThrTyrLeuIleAsnAlaArgLeuSerArgLeuSerGlyMetGly 354
```

Best Local Similarity: 22.73%   Mismatches: 167
Query Match: 6.99%   Indels: 64
DB: 2   Gaps: 13

US-09-270-437D-5 (1-1708) x S41224 (1-396)

Qy   221 ATCCCCTTCGGCTCTCGGTGCCACCCAGTAGTGGTGCCATTATTGGCAAGAGGAGGGG 280
                 ::: ||||||||||::: ::::: ||||||:::||||||||||
Db   40   ValGluLeuArgIleLeuGlnSerLysAsnAlaGlyAlaValIleGlyLysGlyGlyTyr 59

Qy   281 GCCACCATCCGCAACATCACAAAACAGACCCAGTCCAGATAGACGTGCATAGGAAGGAG 340
                 ||||:::   ::: ::: ::: ::::
Db   60   LysAsnIleLysAlaLeuArgThrAspTyrArgAsnAlaSerValSerVal-------Pro 76

Qy   341 AACGCAGGTGCAGAAAAGCCATCAGTGCACTCCACCCCTGAGGGCTGCTCCTCC 400
                 :::: ||| |||::: ::: ::: ||| :::
Db   77   AspSerSerGlyProGluArgIleLeuSerIleSerAlaAspIleGluThrIleGlyGlu 96

Qy   401 GCTTGTTGTAAGATGATCTTGGAGATTATGCATAAAGAGAGCTAAGGACACCAAAACGGCTGAC 460
                 ||| |||||| ::: ::: :::: |||
Db   97   IleLeuLysLysIleIleProThrLeuGluHisPheLysPheLysGlyAsnAspPhe------ 114

Qy   461 GAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAA 520
                 ||||:::|||::: :::||||::: :::||||||
Db   115 AspCysGluLysArgLeuLeuIleHisGlnSerLeuLeuAlaGlyGlyIleIleGlyValLys 134

Qy   521 GGACGGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTG 580
                 ||| ::: ::: :::
Db   135 GlyAlaLysIleLysGluLeuArgGluLysThrGlnThr------ThrIleLysLeuPhe 152

Qy   581 CAAGACCTTACCCTTTACAACCCTGAGAGACCATCATCTGAGAGGGGCCATCAGAGAAT 640
                 ||||::: ::::: :::||||| ::: ::: :::
Db   153 GlnGluCysCysProHisSerThrAspArgValValLeuIleGlyGlyArgProAspArg 172

Qy   641 TGTTGCAGGGCCGAGCAGGAAAATAATGAAGAAAGTTCGGAGGAGGCCTATGAGAATGATGTG 700
                 ||||::: ::: :::: :::::
Db   173 ValValGluCysIleLysValIleLeuLeuAspLeuIleSer-------GluSerProVal 189

Qy   701 GCTGCCATGAGCTCTCACCTGATCCCT 727
                 ||| ::: :::
Db   190 LysGlyArgSerGlnProTyrAspProAsnPheTyrGluThrTyrAspTyrArgGlyPhe 209

Qy   728 --GGCCTGAACCTGGCTGCTGTAGGT 751
                                                                       ||| :::::
Db   210 ThrMetMetPheAspAspArgArgGlyArgProHisGlyPheSerMetHisAlaArgGly 229

Qy   752 CTTTTCCCAGCTTCATCCAGCTCCCGCCGCTCCCCAGCAGCGTTACTGGGGCTGCT 811
                 ||| |||:::||||||||
Db   230 GlyPhe-------------AspArgMetProProGlyPro---------GlyGlyArg 241

Qy   812 CCCTATAGCTTCCTTTATGCAGGCTCCCGAGCAGCAGGAGACGTGTGAGGTGTTTATCCCCGCC 871
                 ||||::: ::: :::
Db   242 -ProMetPro------------------GlnSerArgArgAspTyrAspTyrGlyMetSerProAr 256

Qy   872 CAGGCAGTGGGCGCCATCATGGGCAAGAGAGGGGGCCAGCACATCAAAACAGC---TCTCCCGG 928
                 ||||:::||||||:::||||||:::
Db   256 gArgGlyLeuProProProProGlyLysArgGlyGlyValArgGlyGlyAlaArgAs 276

Qy   929 TTTGGCAGGCCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTCGTATG 988
                 ||||||||| ::: :::
Db   276 nLeuProLeuProPro---Pro---ProProArgGlyGlyAspProArg--------- 291

Qy   989 GTTATCATCACTGACCTCCTTTGTCCCAAGGCTCAGGGAAGAAATCTATGGCAAA 1048
                 ||||::: :::
Db   292 --------------GlyArgProAspHisTyrAspGlyMetGlyTyrArgGlyTyrGlyArg 307

Qy   1049 CTCAAGGAGGAGACTTCTTTGGTCCCAAGGAGGAAGGTGAAGCTGACGAGGTTCCC------CTGAAG 1108
                 ||||||||||::: ::: :::
Db   308 GlySerPheGlyAspIleGlyGlyPro------------------ValIleThrThrGlnValThr 323

Qy   1109 GTGCCAGCATCCAGCAGCTGGCCCGGTCATTGGCAAAGCGGTGAACGAGTTG 1168
                 ::||| ||||||::: :::::
Db   324 IleProLysAspLeuAlaGlySerIleLeuIleGlyLysGlyGlyGlnIleArgGlnIleArgIleIle 343

RESULT 15
T48439
probable RNA-binding protein - Arabidopsis thaliana
N;Alternate names: protein T32M21.30
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T48439
R;Bevan, M.; Terryn, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24487
A;Accession: T48439
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-313 <BEV>
A;Cross-references: EMBL:AL162875
A;Experimental source: cultivar Columbia; BAC clone T32M21
C;Genetics:
A;Map position: 5
A;Introns: 29/1; 108/3; 144/2; 209/1; 229/3; 261/3; 282/2
A;Note: T32M21.30

Alignment Scores:
Pred. No.:   Length: 313
Score: 1.51e-07   Matches: 80
Percent Similarity: 214.00   Conservative: 62
Best Local Similarity: 42.90%   Mismatches: 113
Query Match: 24.17%   Indels: 76
DB: 6.88%   Gaps: 12
                   2

US-09-270-437D-5 (1-1708) x T48439 (1-313)

Qy   185 GCAGCGGGGGCCCCA--- 199
                 |||||||||||::|
Db   8   AlaAlaGlySerProGluGluLeuAlaLysArgSerProGluProHisAspSerSerGlu 27

Qy   200 GCCAAGCAGCAGCAAGTGGACATCCCCCCTTCGGCTCTCGGTGCCACCCAGTATGTGGGT 259
                 |||||||||||::: :::::
Db   28   AlaAspSerAlaGluLysProThrHisIleArgPheLeuValSerAsnAlaAlaAlaGly 47

Qy   260 GCCATTATTGGCAAGAGGAGGGGGCCAGCACATCAAAACAGACCCAGTCCAAG 319
                 |||::||||||:::||||||||||
Db   48   SerValIleGlyLysGlyGlySerThrIleThrGluPheGlnAlaLysSerGlyAlaArg 67

Qy   320 ATAGACGTGCATAGGAAGGAGAAC------GCAGGTGCAGCTGAAAAAGCCATCAGTGTG 373
                 ::: :::: ::: :::
Db   68   IleGlnLeuSerArgAsnGlnLeuProPhePheProGlyThrThrAspArgIleIleMetIle 87

Qy   374 CACTCCACCCCTGAGGGCTGCTCCTCCGTGTAAGATGATCTTGGAGATTATGCATAAAA 433
                 |||::||||||::: :::::
Db   88   SerGlySerIleLysGluValValAsnGlyValLeuGluLeuIleLeuAspLysLeuHisSer 107

Qy   434 GAG------GCTAAGGACACCAAAACGGCTGACGAGGTTCCC---------CTGAAG 475
                 |||::: :::::
Db   108 GluLeuHisAlaGluAsp------------GlyAsnGluValGluProArgArgIleArg 124

Qy   476 ATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGGAACCTGAAG 535
                 ::|||||::: ::: ||||||
Db   125 LeuValIleProAsnSerCysLysLeuIleLeuGluLysGlyGlyGlnIleGlyAlaThrIleLys 144

Qy   1169 CAGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGATGAGAACGAC 1228
                 :::::|||||||:::|||::||||::
Db   344 ArgHisGluSerGlyAlaSerIleLysIle------AspGluProLeuGluGlySerAsp 361

Qy   1229 CAGGTCATCGTGAAAATCATCGACACATTCTATCGCACCAGTCAGATGGCTCAACGGAAGATC 1288
                 |||||:::||| |||::: :::
Db   362 AspArgIleIleThrIleThrGlyThrThrGlnAspGlnIleLeuGlnAsnAlaArgPheLeuLeu 381

Qy   1289 CGAGACATCTCGGCCCCAGGTTAAGCAGCAGCATCAGAAGGGA 1330
                 :::||| :::::||||
Db   382 GlnAsnSerValLysGlnPheSerGluSerGluAspTyrAlaTyrGly 395

```
Qy    536  AAGGTAGAGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCTT  595
                ::::::::::::::     |||   |||||| |||   ||| |||  |||
Db    145  SerPheIleGluGluSerLysAlaGlyIleLeuLysIleSerProLeu---AspAsnThrPhe  163

Qy    596  TACAAC--CCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGGAATTGTTGCAGGGGCC  652
                  ::::|||  :::|||::      |||||::             ||||||
Db    164  TyrGlyLeuSerAspArgLeuValThrLeuSerGlyThrPheGluGlnMetArgAla  183

Qy    653  GAGCAGGAAATAATGAAGAAAGTTCGGGAGGCCCATGAGAATGATGTGGCTGCCATGAGC  712
                |||:::  ||||:||          ||| :::  :::
Db    184  IleAspLeuIleLeuAlaLysLeuThr--------GluAspHisTyrSerGlnAsn  200

Qy    713  TCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAGGTCTTTTCCCAGCTTCATCCAGC  772
                |||
Db    201  ValHis----------------------------------------------  202

Qy    773  GCAGTCCCGCCGCCTCCCAGCACGTTACTGGGGCTGCTCCCATAGCTCCTTTATGCAG  832
                |||        :::  ||||| |||:|:||
Db    203  ---------------SerProTyrSerTyrAlaAlaGlyTyrAsnSerValAsnTyr  216

Qy    833  GCTCCCGAG-----------------------------------CAGGAGATGGTG  853
                |||||||
Db    217  AlaProAsnGlySerGlyGlyLysTyrGlnAsnHisLysGluGluAlaSerThrThrVal  236

Qy    854  CAGGTGTTTATCCCCGCCCAGGCAGTGGGCGCCATCATCCGGCAAGAAGGGGCAGCACATC  913
                :::|||      ::: ||||:::  ||||||::||||::
Db    237  ThrIleGlyValAlaAspGluHisIleGlyLeuValLeuGlyTyrArgGlyGlyAsnIle  256

Qy    914  AAACAGCCTCTCCCGGTTTGCCAGCGCCTCCATCAAGATTGCACCA------------  958
                :::||||| |||  ||||||||||
Db    257  MetGluIleThrGlnMetThrGlyAlaArgIleLysIleSerAspArgGlyAspPheMet  276

Qy    959  CCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGCCAGAGGCCCAA  1018
                ||| ||| |||||||      |||||||||
Db    277  SerGlyThrThrAspArgLysValSer-------IleThrGlyProGlnArgAlaIle  293

Qy    1019 TTCAAGGCTCAGGGAAGAATCTATGCCAAACTC  1051
                :::|||||||::     |||:::
Db    294  GlnGlnAlaGluThrMetIleLysGlnLysVal  304
```

Search completed: July 13, 2004, 12:13:47
Job time : 53 secs

```
              GenCore version 5.1.6
       Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:    July 13, 2004, 11:52:36 ; Search time 19.5 Seconds
                                   (without alignments)
                                   9121.623 Million cell updates/sec

Title:            US-09-270-437D-5
Perfect score:    3110
Sequence:         1 aggagcgctgccgcaccgcc.........atttccttcaggtttaaaa 1708

Scoring table:    BLOSUM62
                  Xgapop  10.0 , Xgapext  0.5
                  Ygapop  10.0 , Ygapext  0.5
                  Fgapop   6.0 , Fgapext  7.0
                  Delop    6.0 , Delext   7.0

Searched:         141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:     283362

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_13072004_121921_9590/app_query.fasta_1.1863
-DB=SwissProt_42 -QPMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN 1 1 24 @runat_13072004_121921_9590 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :        SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	264	8.5	572	1	FUB3_HUMAN	Q96124 homo sapien
2	248	8.0	651	1	FUB1_MOUSE	Q91wj8 mus musculu
3	247.5	8.0	643	1	FUB1_HUMAN	Q96ae4 homo sapien
4	236.5	7.6	492	1	NOA2_HUMAN	Q9unw9 homo sapien
5	235.5	7.6	769	1	ZBP2_CHICK	Q8uvd9 gallus gall
6	234.5	7.5	707	1	FUB2_HUMAN	Q92945 homo sapien
7	231	7.4	474	1	NOA1_RAT	Q80wa4 rattus norv
8	228.5	7.3	721	1	FUB2_RAT	Q99pf5 rattus norv
9	221.5	7.1	356	1	PCB1_HUMAN	Q15365 homo sapien
10	221.5	7.1	493	1	NOA1_MOUSE	Q9jkn6 mus musculu
11	221.5	7.1	510	1	NOA1_HUMAN	P51513 homo sapien
12	218.5	7.0	356	1	PCB1_RABIT	O19048 oryctolagus
13	217.5	7.0	403	1	PCB4_HUMAN	P57723 homo sapien
14	214	6.9	403	1	PCB4_MOUSE	P57724 mus musculu
15	209	6.7	606	1	Y475_ARATH	P58223 arabidopsis
16	209	6.7	1268	1	VGLN_HUMAN	Q00341 homo sapien
17	208.5	6.7	339	1	PCB3_HUMAN	P57721 homo sapien
18	207.5	6.7	339	1	PCB3_MOUSE	P57722 mus musculu
19	203.5	6.5	1270	1	VGLN_CHICK	P81021 gallus gall
20	199	6.4	365	1	PCB2_HUMAN	Q15366 homo sapien
21	197.5	6.4	463	1	ROK_HUMAN	Q07244 homo sapien
22	197.5	6.2	463	1	ROK_RABIT	O19049 oryctolagus
23	194	6.2	362	1	PCB2_MOUSE	Q61990 mus musculu
24	191	6.1	413	1	PBP2_YEAST	P38151 saccharomyc
25	182.5	5.9	1222	1	S160_YEAST	P06105 saccharomyc
26	178	5.7	381	1	YBD2_YEAST	P38199 saccharomyc
27	170.5	5.5	606	1	TDRH_HUMAN	Q9y2w6 homo sapien
28 c	164	5.5	784	1	YAV2_XANCV	P14728 xanthomonas
29	153	4.9	2805	1	MAPA_HUMAN	P78859 homo sapien
30	144.5	4.6	1336	1	W146_HUMAN	Q90j8 homo sapien
31	143	4.6	648	1	GLT0_WHEAT	P10387 triticum ae
32	142	4.6	2142	1	BAT2_HUMAN	P48634 homo sapien
33	141.5	4.5	839	1	GLT5_WHEAT	P10388 triticum ae
34	139.5	4.5	1152	1	MAP4_HUMAN	P27816 homo sapien
35	139.5	4.5	1181	1	HAIR_RAT	P97609 rattus norv
36	138.5	4.5	992	1	EBN6_EBV	P03204 epstein-bar
37	135	4.3	639	1	ZIC5_HUMAN	Q96t25 homo sapien
38	133	4.3	838	1	GLT4_WHEAT	P08489 triticum ae
39	132.5	4.3	1820	1	STR9_HUMAN	Q9p2p6 homo sapien
40	132.5	4.3	2716	1	OSA_DROME	Q8in94 drosophila
41	131	4.2	931	1	MRTA_HUMAN	Q969v6 homo sapien
42	131	4.2	1182	1	HAIR_MOUSE	Q61645 mus musculu
43	131	4.2	5147	1	PCIO_HUMAN	Q9y6v0 homo sapien
44	130.5	4.2	1461	1	IE18_PRVIF	P11675 pseudorabie
45	128.5	4.1	1729	1	TABP_HUMAN	Q9c0c2 homo sapien

ALIGNMENTS

```
RESULT 1
FUB3_HUMAN
ID   FUB3_HUMAN     STANDARD;      PRT;    572 AA.
AC   Q96124; Q92946; Q9BVB6;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Far upstream element binding protein 3 (FUSE binding protein 3).
GN   FUBP3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=97094955; PubMed=8940189;
RA   Davis-Smyth T., Duncan R.C., Zheng T., Michelotti G., Levens D.;
RT   "The far upstream element-binding proteins comprise an ancient family
RT   of single-strand DNA-binding transactivators.";
RL   J. Biol. Chem. 271:31679-31687(1996).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Cervix, and Placenta;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
```

RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May interact with single-stranded DNA from the far-
CC   upstream element (FUSE). May activate gene expression.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=1;
CC   IsoId=Q96I24-1; Sequence=Displayed;
CC   Name=2;
CC   IsoId=Q96I24-2; Sequence=VSP_008323, VSP_008324;
CC   Note=No experimental confirmation available;
CC   -!- TISSUE SPECIFICITY: Detected in a number of cell lines.
CC   -!- SIMILARITY: Contains 4 KH domains.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC   frameshift in position 18.
CC   -!- CAUTION: Ref.2 (AAH01325) sequence differs from that shown due to
CC   a frameshift in position 527.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; U69127; AAC50893.1; ALT_FRAME.
DR   EMBL; BC001325; AAH01325.1; ALT_FRAME.
DR   EMBL; BC007874; AAH07874.1; -.
DR   Genew; HGNC:4005; FUBP3.
DR   MIM; 603536; -.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 3.
DR   SMART; SM00322; KH; 3.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
KW   Transcription regulation; Trans-acting factor; Nuclear protein;
KW   DNA-binding; Repeat; Alternative splicing.
FT   DOMAIN   77   141   KH 1.
FT   DOMAIN   162   228   KH 2.
FT   DOMAIN   253   317   KH 3.
FT   DOMAIN   354   421   KH 4.
FT   VARSPLIC   1   64   MAELVQGQSAPVGMKAEGVDALHRVRQIAAKIDSIPHNN
FT   STPLVDPSVYGVGYGVQKRPLDDGV -> MPPI (in
FT   isoform 2).
FT   /FTId=VSP_008323.
FT   VARSPLIC   298   548   Missing (in isoform 2).
FT   /FTId=VSP_008324.
FT   CONFLICT   72   72   V -> D (IN REF. 1).
FT   SEQUENCE   572 AA; 61640 MW; F1BE223542BC197D CRC64;

Alignment Scores:
Pred. No.:   Length: 572
Score: 1.16e-09   Matches: 129
Percent Similarity: 37.37%   Conservative: 90
Best Local Similarity: 22.01%   Mismatches: 220
Query Match: 8.49%   Indels: 147
DB: 1   Gaps: 22

US-09-270-437D-5 (1-1708) x FUB3_HUMAN (1-572)

```
Qy     23 AGTTTACCCCGGGGAGGAGCCATCATGAAGCTGAAATGGC-----------------CACCAGTTG   67
          ||||||||||||||                                            ||||||||
Db      9 SerAlaProValGly-----MetLysAlaGluGlyPheValAspAlaHisArgVal   26

Qy     68 GAGAACCATGGCCCTGAAGGTCTCTACATCCCGATGAGCAGCAGATAGCACAGGGACCTGAG  127
          |||||:: : :: :::::::          ||||||||||
Db     27 ArgGlnIleAlaAlaLysIleAspSerIleProHisLeuAsnAsnSerThrProLeuVal   46

Qy    128 AATGGGCGCGCCGAGGGGGGTTTGGCTCTCGGGGTCAGCCC---CGCCAGGGCTCACCTGTG  184
          ::: ||||||||||||||                        ::  |||||||
Db     47 AspProSerValTyrGlyValTyrGlyValGlnLysArgProLeuAspAspGlyValGlyAsn   66
```

```
Qy    185 GCAGGCGGGGGCCCCAGCACAGCAGCAAGTGACATCCCCTTCGGCTCCTGGTGCCC  244
          |||||                       |||
Db     67 GlnLeuGlyAlaLeuValHisGlnArgThrVal---IleThrGluGluPheLysValPro   85

Qy    245 ACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGAGGGGCCACATCCCAACATCACAAAA  304
          :::
Db     86 AspLysMetValGlyPheIleIleGlyLysGlyGlyGluGlnIleIleSerArgIleGlnAla  105

Qy    305 CAGACCCAGTCCAAGATAGAGCGTGCATAGGAAGGAGAACGCAGGTGCAGCTGAAAAAGCC  364
          :::::::  ||||||                ::::::::::
Db    106 GluSerGlyCysLysIleGlnIle---AlaSerGluSerGlyIleProGluArgPro  124

Qy    365 ATCAGTGTGCACTCCACCCTGAGGGCTGCTCCACCCTGTTGAAGATGATCTTGGAGATT  424
          |||||||                           ||||||||
Db    125 CysValLeuThrGlyThrProGluSerIleGluGlnIleAlaLysArgLeuGlyIleIle  144

Qy    425 ATG-----------------------------------CATAAAGAGGCTAAGGACACCAAAACCGGCT  457
          :::
Db    145 ValAspArgCysArgAsnGlyProGlyPheHisAsnAspIleAspSerAsnSerThrIle  164

Qy    458 GACGAGGTTCCCCTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAG  517
          |||||                  :: ::|||||
Db    165 GlnGlu---------IleLeuIleProAlaSerLysValGlyValLeuValIleGlyLysArg  180

Qy    518 GAAGGACGGAACCTGGAGAAGTAGACGCAAGATACCGGACAAAATCACCATCCTCCTCG  577
          ||||||||||||                         :: ::
Db    181 GlyGlyGluThrIleLysGlnLeuLysGlnLeuArgThrGlyValLysMetValMet----  198

Qy    578 TTGCAAGACCTTACCCT----------TACAACCCTGAGAGGACCATCACTGTGAAGGGGGCCATC  634
          ::|||||||                                :: :::
Db    199 IleGlnAspGlyProLeuProThrGlyValAlaSpLysProLeuArgIleThrGlyValAspAla  218

Qy    635 GAGAATTGTTGCAGGGCCGAGCAGGAGAAATAATGAAGAAGTTCGGGAGGCCTATGAGAAT  694
          :::  :::  :::::::::  |||||||
Db    219 PheLysValGlnGlnAlaArgGluMetValLeuGluLeuIleAspArgAspSerGlnAla  238

Qy    695 GATGTGGCTGGCTCCATGAGCTCTCACCTGATCCCTGGCCTGACCTGCTGCTGTAGGTCTT  754
          |||                                    :::|||
Db    239 AspPheArgGlyValArgGlyAspPhe-----------------------------  247

Qy    755 TTCCCAGCTTCATCCAGCCAGTCCCGCCTCCCCAGCGTTACTGGGCGTGCTGCTCCC  814
          |||||||||                ||||
Db    248 -----------------------AsnSerArgMetGlyGlyValGlySer-----  255

Qy    815 TATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCGCCCAG  874
          :::::|||                  ::|||
Db    256 --------------------------IleGluValSerValProArgPhe  263

Qy    875 GCAGTGGGGGCCCATCATGCAGGCCTCATCCGCAAGAAGAGGGGCAGCACATCAAACAGTCTCCCCGGTTTGCC  934
          ||||||||            ||||:::  :::
Db    264 AlaValGlyIleLeuValIleGlyLysArgAsnGlyLysMetIleLysTyrSerIleGluAsnAspSerAla  283

Qy    935 AGCGCCTCCATCAAGATTGCACCACCCGAA------ACACCTGACTCCAAAGTTCGTATG  988
          ||||:::::::                    |||
Db    284 GlyValArgIleGlnPheLysProAspAspGlyIleSerProGlu----------ArgAla  300

Qy    989 GTTATCATCACTGACCGCCAGAG----------------------------  1012
          ::: :::
Db    301 AlaGlnValMetGlyProProAspArgCysGlnHisAlaAlaHisIleIleSerGluLeu  320

Qy   1013 ------------GCCCAATTCAAG-----------------------GCTCAGGGAGGAAATC  1039
          :::
Db    321 IleLeuThrAlaGlnGluArgAspGlyPheGlyGlyLeuAlaLeuAlaArgGlyArg---  339

Qy   1040 TATGGCAAACTCAAGAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGCTGAAGCTGAGACC  1099
          ||:::  ::: :::                       ArgAla  300
Db    340 ---GlyArgGlyArgGlyAspTrpSerValGlyValAlaProGlyValGln----GluIle  357

Qy   1100 CACATACGTGTGCCAGCATCAGCAGTGCGGGTCATTGCCAAGGTGAAAAACGGTG  1159
          ||||||||                    |||
Db    358 ThrTyrThrValProAlaAspLysCysSerLeuValIleGlyLysGlyGlyAsnIle  377
```

```
Qy    1160  AACGAGTTGCAGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCTGAT 1219
                  :::   ||   |||   |||   |   |||:::   |||::||   ||||||
Db     378  LysSerIleAsnGlnGlnSerGlyAlaHisValGluLeuGlnArgAsnProProAsn     397

Qy    1220  GAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAA 1279
                                                        :::   |   |||
Db     398  SerAsp---------------------------------------ProAsnLeuArgArg   404

Qy    1280  CGGAAGATCCGAGACATCCTGGCCCAGGTTAAGCAGCAGCATCAGAAG----------- 1327
                  ||||||   :::   ||||:::||
Db     405  PheThrIleArgGlyValProGlnGlnIleGluValAlaArgGlnLeuLeuAspGluLys 424

Qy    1328  -----------------------------------GGACAGAGTAACCAGGCCCAGGCCA 1351
                                                     ||||||||   :::||
Db     425  ValGlyGlyThrAsnLeuGlyLeuAlaProGlyAlaPheGlyGlnSerProPheSerGln--- 443

Qy    1352  CGGAGGAAGTGACCACGCCCCTCCGTCCCTTNGAGTCCAGGACACAACGGGCAGAAAT    1411
                  ||||||   |||                   |||
Db     444  -----------ProProAlaProProHisGlnAsnThrPheProProArgSer        457

Qy    1412  CGAGAGTGTCTCTCCCGGCAGGCCTGAGAATGAGTGGGAATCCG--------------- 1457
                  |||                           ||||||||
Db     458  SerGlyCys-PheProAsnMetAlaAlaLysValAsnGlyAsnProHisSerThrProVa 477

Qy    1458  ---------------------------GGACACNTGGGCGGGCTGTAGATCA        1483
                                             |||   :::||
Db     477  lSerGlyProProAlaPheLeuThrGlnGlyTrpGlySerThrTyrGlnAlaTrpGlnGl 497

Qy    1484  GGTTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGATCNTTCAGCCCCAA 1543
                  :::||                |||   |||||||
Db     497  n-------ProThrGlnGlnValProSerGlnGlnSerGlnProGlnSerGlnProGlnProAs 515

Qy    1544  ACACCCACCCAATTGG 1559
                  :::   ||
Db     515  nTyrSerLysAlaTrp 520
```

RESULT 2

```
FUB1_MOUSE
ID   FUB1_MOUSE              STANDARD;      PRT;    651 AA.
AC   Q91WJ8; Q8C0YB;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Far upstream element binding protein 1 (FUSE binding protein 1) (FBP).
GN   FUBP1 OR D3ERTD330E.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Retina;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Usdin T.B., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
```

```
RN   [2]
RP   SEQUENCE OF 1-286 FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
CC   -!- FUNCTION: Regulates MYC expression by binding to a single-stranded
CC       far-upstream element (FUSE) upstream of the MYC promoter. May act
CC       both as activator and repressor of transcription (By similarity).
CC   -!- SUBUNIT: Interacts with SIAHBP1/FIR and JTV1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q91WJ8-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q91WJ8-2; Sequence=VSP_008322;
CC         Note=No experimental confirmation available;
CC   -!- PTM: Ubiquitinated. This targets the protein for proteasome-
CC       mediated degration (By similarity).
CC   -!- SIMILARITY: Contains 4 KH domains.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; BC014763; AAH14763.1; -.
DR   EMBL; AK029458; BAC26457.1; -.
DR   MGD; MGI:1196294; D3Ertd330e.
DR   MGD; MGI:1924642; 9530027K12Rik.
DR   InterPro; IPR008160; Collagen.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 4.
DR   SMART; SM00322; KH; 4.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
KW   Transcription regulation; Trans-acting factor; Nuclear protein;
KW   DNA-binding; Repeat; Ubl conjugation; Alternative splicing.
FT   DOMAIN        96    160       KH 1.
FT   DOMAIN       181    247       KH 2.
FT   DOMAIN       271    335       KH 3.
FT   DOMAIN       372    439       KH 4.
```

```
FT   DOMAIN          16     21    POLY-GLY.
FT   DOMAIN         345    392    GLY-RICH.
FT   DOMAIN         446    556    PRO-RICH.
FT   VARSPLIC        67     67    D -> DGSWINPSSTTHWEGMPSPFKD (in isoform
FT                                2).
FT                                /FTId=VSP_008322.
FT   CONFLICT        93     93    MISSING (IN REF. 2).
SQ   SEQUENCE       651 AA;  68539 MW;  F4ECA62PD9FA0D5 CRC64;

Alignment Scores:
Pred. No.:              1.3e-08        Length:          651
Score:                  248.00         Matches:         149
Percent Similarity:     35.90%         Conservative:     89
Best Local Similarity:  22.47%         Mismatches:      230
Query Match:            7.97%          Indels:          196
DB:                     1              Gaps:             32

US-09-270-437D-5 (1-1708) x FUB1_MOUSE (1-651)
```

Qy   56 GGCCACCAGTTGGAGAACCATGCCTGCCTGAAGGTCTCCTACATCCCCGATGAGCAGGATAGCA 115
          ||| ||||| ||| ||| :::
Db   19 GlyGlyGlyValValAsnAspAlaPheLysAspAlaLeuGlnArgAlaLeuGlnIleAla 38

Qy   116 CAG-------GGACCTCAGAATGGG-----------CCCCGAGGGGCTTTGGCTCTCGGGGT 160
          ||| ||| || |||| :::: ||||||
Db   39 AlaLysIleGlyGlyAspAlaGlyThrSerLeuAsnSerAsnAspTyrGlyTyrGlyGly 58

Qy   161 CAG---------------------------------CCCCGCCAGGGC 175
          ||| |||
Db   59 GlnLysArgProLeuGluAspGlyAspGlnProAspAlaLysLysValProProGlnAsn 78

Qy   176 TCACCTGTGCAGCCGGCGGCCCCAGCCAGCAGCAAGTGAC---------ATCCCC 226
          ::||||||||||||| || ||| || |||:::   :::
Db   79 AspSerPheGlyAlaGlnLeuProProMetHisGlnGlnSerArgSerValMetThr 98

Qy   227 CTTCGGCTCCTGGTGCCACCCAGTATGTGGGTGCCATTATTGGCCATGGTCCTCCGCTCGT 286
          |||||| |||||||| :: |||||||||||||
Db   99 GluGluTyrLysValProAspGlyMetValGlyPheIleIleGlyArgGlyValGlyGln 118

Qy   287 ATCCGCAAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACGCA 346
          ||| :::: ||| ::   :::::
Db   119 IleSerArgIleGlnGlnGluGlyArgGlyCysLysSerLysGlnIle---AlaProAspSerGly 137

Qy   347 GGTGCAGCTGAAAAAGCCATCAGTGCACTCCACCCCCTGAGGGCTGTCCTCCGCTTGT 406
          |||||||::: || ::   ::||
Db   138 GlyLeuProGluGluArgSerCysMetLeuThrGlyThrProGluSerValGlnSerAlaLys 157

Qy   407 AAGATGATCTTGGAGATTATGCATAAA------GAGGCTCAAGGACACCAAAACGGCTGAC 460
          :::::::: :::||||| ||| |||
Db   158 ArgLeuLeuAspGlnIleValGluGluLysGlyLysGlyIleLysGlyGlyProAlaProGlyPheHisHisGlyHisAsp 177

Qy   461 -------GAGGTTCCCTGAAGATCCTGGCCCATAATAACTTGTAGGGCTGTCTCATT 511
          ||| ::
Db   178 GlyProGlyAsnAlaValAlaGlnLeuMetIleProAlaSerLysAlaGlyLeuValIle 197

Qy   512 GGCAAGGAGGAGGACGAACCTGAAGAGATAGAGCAAGTACCGAGACAAAAATCACCATC 571
          ||||||| :::   ::||||||
Db   198 GlyLysGlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyValLysMetValMet 217

Qy   572 TCCTCGTTGCAAGACCTTACCCTTTACACACCCTGAGAGGACC-------------ATC 616
          ||||||||| || :: ||| :::
Db   218 -----IleGlnAsp---------GlyProGlnAsnThrGlyAlaAspLysProLeu 231

Qy   617 ACTGTGAAGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGAGAAATAATGAAGAAAGTT 676
          ||||||||||| :::||||||
Db   232 ArgIleThrGlyAspProTyrLysValGlnIlnAlaLysGluMetValLeuGluLeuIle 251

Qy   677 CGGGAG------------GCCTATGAGAATGATGTGGCTGCCGCCATGAGCTCCACCTG 721
          ||::: ||::
Db   252 ArgAspGlnGlyGlyPheArgGluValArgAsnGluTyrGly--------- 265

Qy   722 ATCCCTGGCCTGAACCTGCCTGTCTAGGTCTTTTCCCAGGTTCATCCAGCGCCAGTCCCG 781

```
Qy        1663  CACACC------- --------TCAGTGGGAGAAGAAATAAAATTCCTTCA 1698
                    :::|||           :::|||:::   :::    |||:::|||
Db         578  TyrThrLysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaValProAlaProAla 597

Qy        1699  GGT 1701
                    |||
Db         598  Gly 598


RESULT 3
FUB1_HUMAN
ID   FUB1_HUMAN     STANDARD;      PRT;   643 AA.
AC   Q96AE4; Q12828;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Far upstream element binding protein 1 (FUSE binding protein 1) (FBP)
DE   (DNA helicase V) (HDH V).
GN   FUBP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), SEQUENCE OF 191-193; 203-206;
RP   272-279; 284-290; 300-314; 321-328; 394-397; 409-411; 430-438 AND
RP   440-443, AND FUNCTION.
RC   TISSUE=Leukemia;
RX   MEDLINE=94170991; PubMed=8125259;
RA   Duncan R., Bazar L., Michelotti G., Tomonaga T., Krutzsch H.,
RA   Avigan M., Levens D.;
RT   "A sequence-specific, single-strand binding protein activates the far
RT   upstream element of c-myc and defines a new DNA-binding motif.";
RL   Genes Dev. 8:465-480(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 45-64; 133-145; 271-283; 308-321; 379-386; 414-424 AND
RP   430-439.
RX   MEDLINE=21127960; PubMed=11222755;
RA   Vindigni A., Ochem A., Triolo G., Falaschi A.;
RT   "Identification of human DNA helicase V with the far upstream
RT   element-binding protein.";
RL   Nucleic Acids Res. 29:1061-1067(2001).
RN   [4]
RP   PARTIAL SEQUENCE, AND MASS SPECTROSCOPY.
RX   MEDLINE=22166132; PubMed=12176931;
RA   Rappsilber J., Ryder U., Lamond A.I., Mann M.;
RT   "Large-scale proteomic analysis of the human spliceosome.";
RL   Genome Res. 12:1231-1245(2002).
RN   [5]
RP   INTERACTION WITH SIAHBP1.
RX   MEDLINE=20337922; PubMed=10082074;
RA   Liu J., He L., Collins I., Ge H., Libutti D., Li J., Egly J.-M.,
RA   Levens D.;
RT   "The FBP interacting repressor targets TFIIH to inhibit activated
RT   transcription.";
RL   Mol. Cell 5:331-341(2000).
RN   [6]
RP   INTERACTION WITH JTV1, UBIQUITINATION, AND PROTEASOME-MEDIATED
RP   DEGRADATION.
RX   MEDLINE=22716800; PubMed=12819782;
RA   Kim M.J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
RA   Lee S.W., Han J.M., Lee H.-W., Kim S.;
RT   "Downregulation of FUSE-binding protein and c-myc by tRNA synthetase
RT   cofactor p38 is required for lung cell differentiation.";
RL   Nat. Genet. 34:330-336(2003).
RN   [7]
RP   STRUCTURE BY NMR OF 278-447 IN COMPLEX WITH SINGLE STRANDED FUSE DNA.
RX   MEDLINE=21864720; PubMed=11875576;
RA   Braddock D.T., Louis J.M., Baber J.L., Levens D., Clore G.M.;
RT   "Structure and dynamics of KH domains from FBP bound to
RT   single-stranded DNA.";
RL   Nature 415:1051-1056(2002).
CC   -!- FUNCTION: Regulates MYC expression by binding to a single-stranded
CC       far-upstream element (FUSE) upstream of the MYC promoter. May act
CC       both as activator and repressor of transcription.
CC   -!- SUBUNIT: Interacts with SIAHBP1/PIR and JTV1.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q96AE4-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q96AE4-2; Sequence=VSP_008321;
CC         Note=No experimental confirmation available;
CC   -!- PTM: Ubiquitinated. This targets the protein for proteasome-
CC       mediated degradation.
CC   -!- SIMILARITY: Contains 4 KH domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U05040; AAA17976.2; -.
DR   EMBL; BC017247; AAH17247.1; -.
DR   PDB; 1J4W; 06-MAR-02.
DR   Genew; HGNC:4004; FUBP1.
DR   MIM; 603444;
DR   InterPro; IPR008160; Collagen.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 4.
DR   SMART; SM00322; KH; 4.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
KW   Transcription regulation; Trans-acting factor; Nuclear protein;
KW   DNA-binding; Repeat; Ubl conjugation; Alternative splicing;
KW   3D-structure.
FT   DOMAIN        99    163       KH 1.
FT   DOMAIN       184    250       KH 2.
FT   DOMAIN       274    338       KH 3.
FT   DOMAIN       375    442       KH 4.
FT   DOMAIN        13     26       GLY-RICH.
FT   DOMAIN       348    395       GLY-RICH.
FT   DOMAIN       449    559       PRO-RICH.
FT   VARSPLIC     642    643       GQ -> CRFDPASIELAL (in isoform 2).
FT                                 /FTId=VSP_008321.
FT   CONFLICT      96     96       Q -> QS (IN REF. 1).
SQ   SEQUENCE   643 AA;  67473 MW;  086D4EAA0ACF807B CRC64;
```

Alignment Scores:
Pred. No.:   1.4e-08   Length: 643
Score: 247.50   Matches: 140
Percent Similarity: 37.80%   Conservative: 97
Best Local Similarity: 22.33%   Mismatches: 202
Query Match: 7.96%   Indels: 189
DB: 1   Gaps: 32

US-09-270-437D-5 (1-1708) x FUB1_HUMAN (1-643)

Qy   98 CCCGAT--GAGCAGATAGCACAGGGACCTGAGAATGGGCGCGAGGGGGCTTTGGCTCT 154
            ||||| |:|||||||| ||:||||
Db   73 ProAspAlaLysValAla------ProGlnAsn--------AspSerPheGlyThr   87

Qy   155 CGGGGTCACCCCGCCAGGGCTCACCTGTGCAGGCGGGGCCCCAGCCAGCAGCAGACAA 214
            ||||| :::||||||||
Db   88 GlnLeuProProMetHis------------------------GlnGlnGlnArg   97

Qy   215 GTGGACATCCCCTTCGGCTCCTGGTGCCACCCAGTATGTGGGTGCCATTATTGGCCAAG 274
            ||||| :::   :|||| ||||||
Db   98 SerValMetThrGluGluTyrLysValProAspGlyMetValGlyPheIleIleGlyArg 117

Qy   275 GAGGGGGCCACCATCCGCAACATCACAAAACAGACCCCAGTCCAAGATAGACGTGCATAGG 334
            ||||| ::: |||||
Db   118 GlyGlyGlyGlnIleSerArgIleGlyGlnIleGluSerGlyLysLysIle---Ala   136

Qy   335 AAGGAGAAACCGCAGGTGCAGCTGAAAAAGCCATCAGTGCGACTCCACCCCTGAGGGCTGC 394
            :::|| ||| |||||||
Db   137 ProAspSerGlyLysProGluArgSerCysMetLeuThrGlyThrProGluSerVal   156

Qy   395 TCCTCCGCTTGTAAGATGATCTTGGAGATTTATGCATAAA------GAGGCTAAGGACACC 448
            |||||| :::|||| |||
Db   157 GlnSerAlaLysArgLeuLeuAspGlnIleValGluLysGlyArgProAlaProGlyPhe 176

Qy   449 AAAACGGCTGAC--------GAGTTCCCTGAAGATCCTGGCCCATAATAACTTTGTA   499
            ||||| :::||||||
Db   177 HisHisGlyAspGlyProGlyAsnAlaValGlnGlyIleMetIleProAlaSerLysAla 196

Qy   500 GGGGGTCTCATTGGCAAGGAAGGACGGAACCTGAAGAAGTGAGAGCAAGATACCGAGACA 559
            || || ::::|||||||:::
Db   197 GlyLeuValIleLysGlyGlyLysGluThrIleLysGlnIleLeuGlnGluArgAlaGlyVal 216

Qy   560 AAAATCACCATCTCCTCCGTTGCAAGACCTTACCCTTTACAAACCCTGAGAGGACC------ 613
            ||||| |||| |||
Db   217 LysMetValMet------IleGlnAsp---------GlyProGlnAsnThrGlyAla   230

Qy   614 -------ATCACTGTGAAGGGGGCCATCGAGAATTGTTGCAGGGCCGAGCAGGAAATA   664
             ||||||| |||| |||
Db   231 AspLysProLeuArgIleThrGlyAspProThrTyrLysValGlnGlnAlaLysGluMetVal 250

Qy   665 ATGAGAAAGTTCGGGAG----------------GCCTATGAGAATGATGTGGCTGCCATG 709
            :::|| ::::||| |||
Db   251 LeuGluLeuIleArgAspGlnGlyPheArgGluValArgAsnGluThrTyrGly------- 268

Qy   710 AGCTTCACCTGATCCCTGGCTGCTGTGCTGTAGGTGCTTTTCCCAGCTTCATCC   769
            :::||| ||||
Db   269 ----------------------------SerArgIleGlyGly   273

Qy   770 AGCGCAGTCCCGCCGGCCCTCCAGCAGCGGTTACTGGGCCTGCCTCCCTATAGCTTCTTTATG 829
            || |||
Db   269 ------------------------SerArgIleGlyGly------   273

Qy   830 CAGGCTCCCCAGCAGGAGATGGTGCAGTGTTTATCCCCCAGCCAGTGGGCGCCATC   889
            ||||| ::|||
Db   274 -----------AsnGluGlyIleIleAspValProIlePheProArgPheAlaValGlyIleVal 289

Qy   890 ATCCGGCAAGAGGGGCCAGCACATCAAACAGCTCTCCCGGTTTGCCAGCGCCTCCATCAAG 949
            |||||||::: |||:::
Db   290 IleGlyArgAsnGlyGluMetIleLysLysIleGlnAsnAspAlaGlyValArgIleGln   309

Qy   950 ATTGCACCACCCGGA-------ACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGA 1003
            || ::: |||:::
Db   310 PheLysProAspAspGlyThrThrProGlu---------ArgIleAlaGlnIleThrGly   326

Qy 1004 CCGGCAGAG---GCCCAATTCAAGGCTCAG-----   1030
            |||||| |||:::
Db   327 ProProAspArgCysGlnHisAlaAlaAlaGluIleIleThrAspLeuLeuArgSerValGln 346

Qy 1031 --------------------------------GGAAGAATCATGGCAAACTCAAGGAGGAG 1060
                                                                :::
Db   347 AlaGlyAlaAsnProGlyProGlyProGlyArgGlyArgGlyArgGlyGlnGlyAsn   366

Qy 1061 AACTTCTTTGGTCCAAGGAGGAGAAGTGAAGCTGGAGACCCACCTACGTGTGCCAGCATCA 1120
            ||||||| ||||||| ||
Db   367 TrpAsnMetGlyProProGlyGlyLeuGln---GluPheAsnPheValProThrGly   385

Qy 1121 GCAGCTGGCCGGGTCATTGGCAAAGTGCAAAACGGTGAACGAGTTGCAGAATTCGACG   1180
            |||||::: ||||||||:::|||||
Db   386 LysThrGlyLeuIleLeuGlyLysGlyGlyGluThrIleLysSerIleSerGlnLnSer   405

Qy 1181 GCAGCTGAGGTGGTAGTACCAAGAGACAGACCCTGATGAGAACGACCAGGTC---ATC 1237
            |||||| ::: ||||:::|||
Db   406 GlyAlaAlaArgIleGluLeuGlnArgAsnProProAsnAlaAspProAsnMetLysLeu 425

Qy 1238 GTGAAAATCATCGGACAT-----------------TTCTATGCCAGTCAGATGGCTCAACGGAAG 1285
            ||||||| :::||||||||
Db   426 PheThrIleArgGlyThrProGlnIleGlnIleAspTyrAlaArgGlnLeuIleGluGluLys 445

Qy 1286 AT-------CCAGACAATCCT---GGCCCAGGTTAAGCAGCAGCATCAGAGGGGACAGAG 1335
             || :::|||||| |||:::
Db   446 IleGlyIleGlyProValAsnProLeuGlyProProHisGlyProHisGlyProHisGlyVal--- 464

Qy 1336 TAACCAGGCCCCAGGCACGGA------------GGAAGTGACCAGCCCCTC-------- 1373
            ||||||| ||||||||
Db   465 ---ProGlyPro--HisGlyProProGlyProGlyProThrProMetGlyProT   483

Qy 1374 -------CCTGTCCCTTNGAGTCCAGGA   1394
                   ||||||||:::
Db   483 yrAsnProAlaProTyrAsnProGlyProProGlyProAlaProHisGlyProAlaProAlaP 503

Qy 1395 ----------------------------CAACAACGGGCAGAAATCG   1413
                                        |||||||:::
Db   503 roTyrAlaProGlnGlyTrpGlyAsnAlaTyrProHisProGlnGlnInAlaProProA   523

Qy 1414 AGAGTGTGTCTCCCCGGCAGGCCTGAGAATCAGTGGGAATNCCGGACACNTGGGCCCGGG 1473
            ||| ||||:::
Db   523 sp------------ProAlaAlaLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAla 538

Qy 1474 CTTGTAGATCAGGTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCCTGATCTN 1533
            |||||| :::|||||
Db   539 Tyr--------TyrAlaHisTyrTyrGlnGlnAla-   548

Qy 1534 TCAGCCCCAAACACCCACCCAATTGGC---CCAACACTGTNTGCCCCTCGGGGTGTCAGA 1590
            :::|| ||||||
Db   549 GlnProProProAlaAlaProThrAlaProThrHisThrArgGlnThrAsnGlyLnGlyGly 568

Qy 1591 AATTNTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAAGAAGCTTCCAGGCCCCACC 1650
                                     |||:::|||
Db   569 Asp-------------------------------GlnGlnAsnProAlaProAla   576

Qy 1651 AAGAGGGTGGATCACACC------------------------TCAGTGGGAAGAAAAATA 1686
            ||||||||||:::||||||
Db   577 GlyGlnValAspTyrThrLysAlaTrpGluTyrTyrLysLysMetGlyLnAlaVal   596

Qy 1687 AAATTCCTTCAGGT 1701
            |||:::
Db   597 ProAlaProThrGly 601

RESULT 4
NOA2_HUMAN
ID   NOA2_HUMAN   STANDARD;   PRT;   492 AA.
AC   Q9UNW9; O43267; Q9UEA1;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   RNA-binding protein Nova-2 (Neuro-oncological ventral antigen 2)

```
DE   (Astrocytic NOVA1-like RNA-binding protein).
GN   NOVA2 OR ANOVA OR NOVA3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20197319; PubMed=10735272;
RA   Ueki K., Ramaswamy S., Billings S.J., Mohrenweiser H.W., Louis D.N.;
RT   "ANOVA, a putative astrocytic RNA binding protein gene that maps to
RT   chromosome 19q13.3.";
RL   Neurogenetics 1:31-36(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99007301; PubMed=9789075;
RA   Yang Y.Y., Yin G.L., Darnell R.B.;
RT   "The neuronal RNA-binding protein Nova-2 is implicated as the
RT   autoantigen targeted in POMA patients with dementia.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:13254-13259(1998).
RN   [3]
RP   SEQUENCE OF 29-492 FROM N.A.
RA   Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA   Burkhart-Schultz K.J., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA   Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA   Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA   Attix C., Andreise T., Trankheim M., Amico-Keller G., Coefield J.,
RA   Duarte S., Lucas S., Bruce R., Thomas P., Quan D., Kronmiller B.,
RA   Arellano A., Brower A., Sanders C., Ow D., Nolan M., Trong S.,
RA   Kobayashi A., Olsen A.S., Carrano A.V.;
RT   "Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
RT   D19S412.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 405-480.
RX   MEDLINE=99148126; PubMed=10368286;
RA   Lewis H.A., Chen H., Edo C., Buckanovich R.J., Yang Y.Y.-L.,
RA   Musunuru K., Zhong R., Darnell R.B., Burley S.K.;
RT   "Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding
RT   domains.";
RL   Structure 7:191-203(1999).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 405-491 IN COMPLEX WITH RNA.
RX   MEDLINE=20139984; PubMed=10676814;
RA   Lewis H.A., Musunuru K., Jensen K.B., Edo C., Chen H., Darnell R.B.,
RA   Burley S.K.;
RT   "Sequence-specific RNA binding by a Nova KH domain: implications for
RT   paraneoplastic disease and the fragile X syndrome.";
RL   Cell 100:323-332(2000).
CC   -!- FUNCTION: May regulate RNA splicing or metabolism in a specific
CC       subset of developing neurons (By similarity). Binds single strand
CC       RNA.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- TISSUE SPECIFICITY: Brain. Expression restricted to astrocytes.
CC   -!- DOMAIN: The third KH domain (KH3) recognizes specifically 5'-
CC       UCAY-3'.
CC   -!- DISEASE: Autoantigen in the paraneoplastic opsoclonus myoclonus
CC       ataxia (POMA), a paraneoplastic neurological syndrome/disorder
CC       (PNS/D) associated with breast cancer, fallopian cancer, and SCLCa
CC       and characterized primarily by loss of inhibitory control of motor
CC       neurons in the spinal cord and brainstem. Recognized by the IgG
CC       autoantibody ANNA-2 (also called anti-Ri).
CC   -!- SIMILARITY: Contains 3 KH domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U70477; AAB88661.1; ALT_INIT.
DR   EMBL; AF083898; AAC72355.1; -.
DR   EMBL; AC006540; AAD13116.1; -.
DR   PDB; 1DT4; 18-FEB-00.
DR   PDB; 1EC6; 07-APR-00.
DR   Genew; HGNC:7887; NOVA2.
DR   MIM; 601991; -.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 3.
DR   SMART; SM00322; KH; 3.
DR   PROSITE; PS50084; KH_TYPE_1; 3.
KW   Antigen; Nuclear protein; RNA-binding; Repeat; 3D-structure.
FT   DOMAIN        10     26       BIPARTITE NUCLEAR LOCALIZATION SIGNAL
FT                                 (POTENTIAL).
FT   DOMAIN        32     99       KH 1.
FT   DOMAIN       130    196       KH 2.
FT   DOMAIN       232    405       ALA-RICH.
FT   DOMAIN       332    393       GLY-RICH.
FT   DOMAIN       406    473       KH 3.
FT   CONFLICT     247    247       A -> R (IN REF. 1).
FT   CONFLICT     265    267       PAA -> TAT (IN REF. 1).
SQ   SEQUENCE   492 AA;  49008 MW;  41B63EAF68992565 CRC64;
```

Alignment Scores:
```
Pred. No.:            6.8e-08           Length:           492
Score:               236.50            Matches:          118
Percent Similarity:   37.89%           Conservative:      62
Best Local Similarity: 24.84%          Mismatches:       185
Query Match:           7.60%           Indels:           111
DB:                    1                Gaps:              15
```

US-09-270-437d-5 (1-1708) x NOA2_HUMAN (1-492)

```
Qy   152 TCTCGGGGTCAGCCCCGCCCAGGGCTCACCT------GTGGCAGCGGGGGCCCCAGCGCAAG 205
             ||||| | ::|||||  ::|     :||||                            |||
Db     8 SerArgLysArgProLeuGluThrProGluValCysThrLysArgSerAsnThr        27

Qy   206 CAGCAGCAAGTGGACATCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCCATT   265
             | ||||||||   :  ::|    ||  ||::::|||||||      ::|  |||||||||
Db    28 GlyGluGluGlyGlyTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlySerIle   47

Qy   266 ATTGGCAAGGAGGGGGCCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGAC 325
             |||||||   ::: |   || | || : :  ||| ||| :::|||  :::||  ||||
Db    48 IleGlyLysGlyGlyGlnThrIleValGlnLeuGlnLysGluThrGlyAlaThrIleLys  67

Qy   326 GTGCATAGGAAGGAGAAC------GCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCC  379
             ::: :: :::: :::::        |||||||||||||||  :||:     |||
Db    68 LeuSerLysSerLysSerAspPheTyrProGlyThrThrGluArgValCysLeuValGlnGly  87

Qy   380 ACCCCTGAGGGCTGCTCCTCGCTTGTAAGATGATCTTGAAGATTATGCATAAAGGAGGCT 439
             |||||||||||   ||   |||   :|||   |:||  :::|||  ||||
Db    88 ThrAlaGluAlaLeuAsnAlaValHisSerPheIleAlaGluLysVal---ArgGluIle 106

Qy   440 AAGGCACCCAAAACGGCTGACGAGGTT-----------                       466
             ||||                 |||||||
Db   107 ProGlnAlaMetThrLysProGluValAlaAsnIleLeuGlnProGlnThrThrMetAsn 126

Qy   467 ------------------CCCCTGAAGATCCTGGCCCCATAATAACTTTGTAGGGCGTCTCATT 511
                               |||::||    :|||  ::|||    ||  |||
Db   127 ProAspArgAlaLysGlnAlaLysLeuIleValProAsnSerThrAlaGlyLeuIleIle 146

Qy   512 GGCAAGGAAGGACGGAACCTGAAGAAGGTAGAGGAAGATACCGAAGCAAAAATCACCATC 571
             |||||||| |||          |||   |||  ||:      |:: :::     ::: :||
Db   147 GlyLysGlyGlyGlnThrValLysAlaValMetGluGlnSerGlyAlaTrpValGlnLeu 166

Qy   572 TCC--------------------TCGTTGCAAGACCTTACCCTTTACACCCTGAGAGGACC 613
             |||                     ::|||||||
Db   167 SerGlnLysProGluGlyIleAsnLeuGln----------------------GluArgVal 179

Qy   614 ATCACTGTGAAGGGGGCCATCGAGAATTGTTGCAGGGCCAGCAGGAAAATAATGAAGAAA 673
             ||:||||  |::  :|| |:::|||   ||||::|  ::|||
Db   180 ValThrValSerGlyGluProGlnValHisLysAlaValSerAlaIleValGlnLys   199
```

```
Qy    674 GTTCGGGAG-----------------------------------GCCTATGAGAAT-------- 694
          ||||:|||                                    ::|||||||
Db    200 VisGlnGluAspProGlnSerSerSerCysLeuAsnIleSerTyrAlaAlaAsnValAlaAlaGly 219

Qy    695 -----------------------------------------------------GATGTG------ 700
                                                               ||||||
Db    220 ProValAlaAsnSerAsnProThrGlySerProTyrAlaSerProAlaAspValLeuPro 239

Qy    701 -----------GCTGCCATGAGCTCTCACCTGATCCCTGGCCTGAACCTGGCT-------- 742
                     |||||||||  :::|||||||  ||||::|||
Db    240 AlaAlaAlaAlaAlaSerAlaAlaAlaAlaSerGlyLeuLeuGlyLeuAlaGlyLeuAla 259

Qy    743 GCTGTAGGTCTTTTCCCAGCTTC-----------------------ATCCAGCGC 774
          |||||||                                       :::
Db    260 GlyValGlyAlaPheProAlaAlaLeuProAlaPheSerGlyThrAspLeuLeuAlaIle 279

Qy    775 AGTCCCCGCCGCCTCCCAGCAGGTTACTGGGCGTCTCCCTATAGCTCCTTATGCAGGC 834
          |||  |||    :::                    ||| |||
Db    280 SerThrAlaLeuAsnThrLeuAlaSerTyrGlyTyrAsnThrAsnSerLeuGlyLeuGly 299

Qy    835 TCCCAGCAGGAGATGGTGCAGGTGTTTATCCCCGCCAGGCAGTGGGCGCCATCATCG 894
          :::                                     ::: :::
Db    300 LeuAsnSerAlaAlaAlaSerGlyValValLeuAlaAlaValAlaValAlaAlaAsnProAla 319

Qy    895 CAAGAGAGGGGCAGCACATCAAACAGCTCTCCCGGTTGCCAGCGCGCCTCCATCAAGATTGC 954
          |||||||:::|:::
Db    320 AlaAlaAlaAlaAlaAsnLeuLeuAla-SerTyrAlaGlyAlaGlyLeuAlaGlyAlaGlyProGl 339

Qy    955 A----------CCACCCGAAACACCTGACTCCAAGTTCGTATGGTTATCATCAC 999
          |||||      ||||||                          :::
Db    339 aGlyGlyAlaGlyAlaProProProProProProGlyAlaLeuGlySerPheAlaLeuAlaLeuAsnL 359

Qy   1000 T-----------------------------------------------------GG 1002
          |||
Db    359 aAlaAlaAlaAlaAsnGlyTyrLeuGlyAlaGlyAlaGlyGlyAlaGlyGlyAlaGlyGlyGlyGl 379

Qy   1003 ACCGGCCAGAGCGCCCCAATTCAAGGCTCAGGGAAGAGATCTATGGCAAACTCAAGGAGGAGAA 1062
          |||||||||||||||||||||||||
Db    379 yProLeuValAlaAlaAlaAlaAlaAlaAlaGlyAlaAlaAlaGlyGlyAlaGlyProPheLeuThrAlaGluLeuV 399

Qy   1063 CTTCTTTGGTCTCCAAGGAGGAAGTGAAGTGAGACCCACATACGTGTGCCAGCATCAGC 1122
          ||||||  ||||||||||
Db    399 sLeu-----AlaAlaGluSerAlaLysGluLeuGluLeuAlaValProGlyAsnLe 417

Qy   1123 AGCTCGGCCGGGTCATTGGCAAAGGCGTCAGGGAAAAAACCGTGAACGAGTTGCAGAATTTGACCGC 1182
          |||||||||||||  |||:::    |||||||||
Db    417 uValGlyAlaIleLeuGlyLysGlyGlyLysThrLeuValGluThrArgGlnLeuThrGl 437

Qy   1183 AGCTGAGGTGGTAGTACCAAGA------GACCAGACCCCTGATGAGAACGACCAGGTCAT 1236
          |||||  :::  :::          :::
Db    437 yAlaArgIleGlnIleSerIysLysGlyGluPheLeuProGlyThrArgAsnArg---Ar 456

Qy   1237 CGTGAAAATCATCCGACATTTCATGCCAGTTCAGATGGCTCAA 1279
          |||||||||:::|||||||
Db    456 gValThrIleThrGlySerProAlaAlaThrGlnAlaAlaAlaGln 470
```

RESULT 5
```
ZBP2_CHICK
ID   ZBP2_CHICK     STANDARD;      PRT;      769 AA.
AC   Q8UVD9;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Zipcode-binding protein 2.
GN   ZBP2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 104-119; 242-267; 679-685; 705-711 AND
RP   743-766, FUNCTION, DEVELOPMENTAL STAGE, AND SUBCELLULAR LOCATION.
RX   MEDLINE=21640447; PubMed=11781334;
RA   Gu W., Pan F., Zhang H., Bassell G.J., Singer R.H.;
RT   "A predominantly nuclear protein affecting cytoplasmic localization of
RT   beta-actin mRNA in fibroblasts and neurons.";
RL   J. Cell Biol. 156:41-52(2002).
CC   -!- FUNCTION: Binds to a 54-nucleotide localization signal (the
CC       zipcode) found in the 3' untranslated region of beta-actin mRNA
CC       and may play a role in its localization.
CC   -!- SUBCELLULAR LOCATION: Predominantly nuclear. Also detected in the
CC       cytoplasm.
CC   -!- DEVELOPMENTAL STAGE: The highest expression is found in 6-d
CC       embryos, is reduced to 30% before hatching and remains stable
CC       thereafter.
CC   -!- SIMILARITY: Contains 4 KH domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF461020; AAL66365.1; -.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0003730; F:mRNA 3' UTR binding; IDA.
DR   GO; GO:0008298; P:mRNA localization, intracellular; NAS.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 4.
DR   SMART; SM00322; KH; 4.
KW   Transport; mRNA transport; Nuclear protein; RNA-binding; Repeat.
FT   DOMAIN      208    272       KH 1.
FT   DOMAIN      296    362       KH 2.
FT   DOMAIN      385    449       KH 3.
FT   DOMAIN      486    553       KH 4.
SQ   SEQUENCE    769 AA;   80644 MW;  875A6C83529969EB CRC64;
```

Alignment Scores:
```
Pred. No.:                            8.76e-08        Length:        769
Score:                                235.50          Matches:       134
Percent Similarity:                   32.59%          Conservative:  70
Best Local Similarity:                21.41%          Mismatches:    189
Query Match:                          7.57%           Indels:        233
DB:                                   1               Gaps:          26
```

US-09-270-437D-5 (1-1708) x ZBP2_CHICK (1-769)

```
Qy    10 GCCGGCACCGGCCGCCCCAGTTTACCCCGGGGAGCCATCATGAAGCTGAATGGCACCACCAGTTGGA 69
         |||||||||||||||||||||||:::        :::
Db    155 AlaAlaAlaProProGluArgSer----------------------GlyProValGly 165

Qy    70 GAACCATGCCCTGAAGTCTCCTACATCCCCGATGAGCAGATAGCACAGGACCTGAGAA 129
         ::||||:::
Db    166 AspProProGlyPro----------------ProArg---------------ProArg 172

Qy   130 TGGGCCGCCGAGGGGGTCTTTGGCTCTGGGGTCAGCCCGCCAGGGCTCACCTGTCGGC--- 186
         ||||||                  |||        |||||
Db    173 ---AlaGluAlaArgGly-----------ArgArgProProAlaLeuThrGlyGlyAla 187

Qy   187 ------------AGCGGGGGCCCCAGCCAAGCAGCAGCAAGTGACAATGCCCCCCTTCGG--- 232
                    ||| :::              |||
Db    188 LeuProSerAlaAlaLeuProProGlnLeuGlyProMetHisProProPro-ArgSerTh 207

Qy   233 ------------CTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGA 276
                     :::              |||||
Db    207 rThrValThrGluTyrArgValProAspGlyMetValGlyLeuIleGlyArgGl 227

Qy   277 GGGGGCCACCATCCGCAACATCCAACATCAGAACCCAGTCCAAGATAGACGTGCATAGGAA 336
         ||||            :::        :::
```

```
Db  227 yGlyGluGlnIleAsnLysIleGlnGlnAspSerGlyCysLysValGlnIle---SerPr 246

Qy  337 GGAGACGCAGGTGCAGCTGAAAAGCCATCAGTGTGCACTCCACCCTGAGGGCTGCTC 396
        :::::::     |||   ::::::::::|||:
Db  246 oAspSerGlyGlyLeuProGluArgGlySerValSerLeuThrGlySerProGluAlaValGl 266

Qy  397 CTCCGCTTGTGAAGATGATCTTGGAGATTATG------------------ 427
        |||   ::|||:::
Db  266 nLysAlaLysLeuMetLeuAspAspIleValSerArgArgGlyArgGlyProGlyGl 286

Qy  428 ---CATAAAGAGGCTAAG--GACACCAAAACGGCTGACGAGGTTCCCGTGAGGATCCT 480
        |||            :::
Db  286 nPheHisAspTyrAlaAsnGlyGlnAsnGlyThrValGlnGlu---------IleMe 302

Qy  481 GGCCCATAATAACTTGTAGGGCGTCTCATTGGCAAGGAAGGACCGGAACCTGAAGAAGGT 540
        ||| :::|||:::::
Db  302 tIleProAlaGlyLysGlyLysLeuValIleGlyLysGlyGlyGluThrIleLysLysLeuGlnuLys 322

Qy  541 AGAGCAAGATACCGAGACAACAAAATCACCATC------TCCTCGTTGCAAGACCTTACCCT 594
        ::::::|||
Db  322 uGlnGluArgAlaGlyValLysValLysMetIlePheIleGlnAspGlySerGlnAsnThrAsnVa 342

Qy  595 TTACAACCCTGAGAGACCATCACTGTGAAGGGGCCATCAGAGAATTGTTGCAGGGCCGA 654
        :::::::
Db  342 lAspLysProLeuArgIleIleGlyIleLeuGlyAspProTyrLysValGlnAlaCysGluMetVa 362

Qy  655 GCAGGAAATAATGAGAGAAGTTCGGGAGGCCTATGAGAATGATGTGGCTGCCATGAGCTC 714
        ::|||:::::::   :::
Db  362 lMetAspSerIleLeuLeuArgGluArgGlyAspPheGlyGlyPheGlyAspArgAsnGluTyrGlySe 382

Qy  715 TCACCTGATCCCTGGCCTGAACCTGGCTGCGTAGGGTGTTTTCCCAGCTTCATCCAGCGC 774
        |||:::
Db  382 rArgIleGlyIleGlyIleAsp---------------------------- 389

Qy  775 AGTCCCGCCGGCCTCCCCAGCAGGTTACTGGGCGTGCTGCTCCCCTATAGCTCCTTATGCAGGC 834
        |||||
Db  390 -ValPro---------------------------------------------- 391

Qy  835 TCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCCAGCAGTGGGCGCCATCATCGG 894
        :::|||||||
Db  392 ------------------------ValProArgHisSerValGlyValValIleGl 402

Qy  895 CAAGAAGGGGCAGCACACTCAAACAGCTCTCCGGTTGCCAGCGCGCCTCCATC---------- 946
        |::: |||::: ||||||::::::
Db  402 yArgSerGlyGluMetIleLysLeuTyrIleLeuGlnAsnAspAlaGlyValArgIleGlnInPheLy 422

Qy  947 ---------------AAGATTCA---------CCACCCGAAACACC 969
        |||||||
Db  422 sGlnAspAspProGlyThrGlyProProGlyHisMetProProGlyGluArgProGluProGluArgCy 442

Qy  970 TGACTCCAAAGTTCGTATGGT--------ATCATCACTGACACCGCCAGA 1011
        :::::::|||||||
Db  442 sGluHisAlaAlaAlaArgIleIleIleAsnAspLeuLeuGlnSerLeuArgSerGlyProPro-- 461

Qy  1012 GGGCCCAATT------------CAAGGCTCAGGGAAGAATCTA 1041
        ||||||            :::||| ||||:::|||
Db  462 GlyProProGlyHisGlyMetProProGlyProGlyArgGlyArgGlyGlnGlyIle 481

Qy  1042 TGGCAAACTCAAGGAGGAGGAAACTTCTTTGGTCCCAAGGAGGAAGGTGAAGCTGGAGACCCA 1101
        |||||||::: :::
Db  482 TrpGlyTyrProGlyProGlyGlyGlyMetThrPhe------------------ 491

Qy  1102 CATACGTGTGCCAGCATCAGCGCTGGGGCTGTGGGSGTLyLeuValIleGlyArgGlyGluAsnVallLy 510
        |||||||::::::   :::||
Db  492 ---Ser-IleProThrHisLysCysGlyArgCysTrpArg 510

Qy  1162 CGAGTTGCAGATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAGACCCCT---- 1216
        :::||| ||||:::
Db  510 sAlaIleAsnGlnGlnArgGlyAlaPheValGluIleSerArgGlnLeuProProAsnGl 530

Qy  1217 -GATGAGAACGACCAGTCATCGTGAAAATCATCGGACATTCTATGCCAGTCAGATGGC 1275
        ::::::
Db  530 yAspProAsnPheLysLeuPheIle--------------------------- 538
```

```
Qy  1276 TCAACGGAGATCCGAGACATCCTGGCCCAGGTTAAGCAGCAGCAGCATCAGAAGGGACAGAG 1335
        |||||                ::::::
Db  539 --------IleArgGlySerProGlnGlnIleGluHisAlaAlaLysGln---------- 551

Qy  1336 TAACCAGGCCCAGGCACGGAGGAAGTGACCAGCCCCTCCCGTCCTTNGAGTCCAGGAC 1395
        ||| ||||||
Db  552 ---------ProIleGluGlyLysIleGluGlyLyProLeuCysPro------------ 563

Qy  1396 AACAACGGGCAGAAATCCAGAGTGTGCTCTCCCCGGCCAGGCCTGAGAATGAGTGGGAATC 1455
        |||:::
Db  563 ---------------------------------------------- 563

Qy  1456 CGGGACACNTGGGCGGCGGCTGTAGATCCAGTTCAGTTGTTGCCCACTTGATTGAGAGAAAGATGTTCCA 1515
        ||||
Db  564 ------ValGlyProGlyProGlyProGlyProGlyPro------------------ 574

Qy  1516 GTGAGGAACCCTGATCTNTCAGCCCCAAACACCCACCACTTGGCCCAACACTGTNTGCC 1575
        :::
Db  575 --------------------------------------GlyPr 576

Qy  1576 CCTCGGGGGTGTCAGAAATNTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAAGAAGC 1635
        ||| :::::::
Db  576 oAlaGly---------------------ProMetGlyProPheAsnProGlyProPheAsnGlnGl 591

Qy  1636 T---------------------------------CTCCAGGCCCCACCAAGAGGGTGG-----ATCACACC 1668
        |||:::||:|||
Db  591 yProProGlyProGlyProProHisGlnTyrProProGlnGlyTrpProGlyTrppGlyAsnThrTyrPr 611

Qy  1669 TCAGTGGGAA 1678
        |||:::
Db  611 oGlnTrpGln 614
```

RESULT 6

FUB2_HUMAN

```
ID   FUB2_HUMAN   STANDARD;   PRT;   707 AA.
AC   Q92945; O00301; Q9UNT5; Q9UQH5;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Far upstream element binding protein 2 (FUSE binding protein 2) (KH
DE   type splicing regulatory protein) (KSRP) (p75).
GN   FUB2P OR KHSRP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 72-85; 123-128; 267-281; 283-291;
RP   348-359; 472-486; 488-492; 620-625; 627-644 AND 646-651, AND FUNCTION.
RC   TISSUE=Neuroblastoma, and Retinoblastoma;
RX   MEDLINE=97282621; PubMed=9136930;
RA   Min H., Turck C.W., Nikolic J.M., Black D.L.;
RT   "A new regulatory protein, KSRP, mediates exon inclusion through an
RT   intronic splicing enhancer.";
RL   Genes Dev. 11:1023-1036(1997).
RN   [2]
RP   SEQUENCE OF 1-115 AND 570-707 FROM N.A.
RX   MEDLINE=99189245; PubMed=10087204;
RA   Ring H.Z., Vameghi-Meyers V., Nikolic J.M., Min H., Black D.L.,
RA   Francke U.;
RT   "Mapping of the KHSRP gene to a region of conserved synteny on human
RT   chromosome 19p13.3 and mouse chromosome 17.";
RL   Genomics 56:350-352(1999).
RN   [3]
RP   SEQUENCE OF 60-707 FROM N.A., AND FUNCTION.
RC   TISSUE=B-cell lymphoma, and Skeletal muscle;
RX   MEDLINE=97094955; PubMed=8940189;
RA   Davis-Smyth T., Duncan R.C., Zheng T., Michelotti G., Levens D.;
RT   "The far upstream element-binding proteins comprise an ancient family
RT   of single-strand DNA-binding transactivators.";
RL   J. Biol. Chem. 271:31679-31687(1996).
RN   [4]
```

```
RP   FUNCTION, AND INTERACTION WITH PTBP1; PTBP2 AND HNRPH1.
RX   MEDLINE=20459250; PubMed=11003644;
RA   Markovtsov V., Nikolic J.M., Goldman J.A., Turck C.W., Chou M.-Y.,
RA   Black D.L.;
RT   "Cooperative assembly of an hnRNP complex induced by a
RT   tissue-specific homolog of polypyrimidine tract binding protein.";
RL   Mol. Cell. Biol. 20:7463-7479(2000).
CC   -!- FUNCTION: Binds to the dendritic targeting element and may play a
CC       role in mRNA trafficking (By similarity). Part of a ternary
CC       complex that binds to the downstream control sequence (DCS) of the
CC       pre-mRNA. Mediates exon inclusion in transcripts that are subject
CC       to tissue-specific alternative splicing. May interact with single-
CC       stranded DNA from the far-upstream element (FUSE). May activate
CC       gene expression.
CC   -!- SUBUNIT: Part of a ternary complex containing FUBP2, PTBP1, PTBP2
CC       and HNRPH1.
CC   -!- SUBCELLULAR LOCATION: Nuclear. A small proportion is also found in
CC       the cytoplasm of neuronal cell bodies and dendrites (By
CC       similarity).
CC   -!- TISSUE SPECIFICITY: Detected in neural and non-neural cell lines.
CC   -!- SIMILARITY: Contains 4 KH domains.
CC   -!- CAUTION: Ref.3 sequence differs from that shown due to numerous
CC       frameshifts.
```

```
DR   EMBL; U94832; AAB53222.1; -.
DR   EMBL; AF093747; AAD29861.1; -.
DR   EMBL; AF093745; AAD29861.1; JOINED.
DR   EMBL; AF093748; AAD29862.1; -.
DR   EMBL; AF093748; AAD29862.1; ALT_FRAME.
DR   EMBL; U69126; AAC50892.1; ALT_FRAME.
DR   Genew; HGNC:6316; KHSRP.
DR   MIM; 603445; -.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 4.
DR   SMART; SM00322; KH; 4.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
KW   Transport; mRNA transport; mRNA processing; mRNA splicing;
KW   Transcription regulation; Trans-acting factor; Nuclear protein;
KW   DNA-binding; RNA-binding; Repeat.
FT   DOMAIN      144    208    KH 1.
FT   DOMAIN      233    299    KH 2.
FT   DOMAIN      322    386    KH 3.
FT   DOMAIN      422    489    KH 4.
FT   DOMAIN      569    682    4 X 12 AA IMPERFECT REPEATS.
FT   REPEAT      569    580    1.
FT   REPEAT      615    626    2.
FT   REPEAT      641    652    3.
FT   REPEAT      671    682    4.
FT   DOMAIN        7     67    GLY/PRO-RICH.
FT   DOMAIN       68    495    GLY-RICH.
FT   DOMAIN      496    610    ALA/GLY/PRO-RICH.
FT   CONFLICT     96     96    G -> V (IN REF. 2).
FT   CONFLICT    692    692    V -> G (IN REF. 2).
FT   CONFLICT    694    694    G -> GG (IN REF. 2).
FT   CONFLICT    707    707    Q -> A (IN REF. 2).
SQ   SEQUENCE   707 AA;  72709 MW;  E07588DE43BCA8B6 CRC64;
```

Alignment Scores:

Pred. No.:	9.97e-08	Length: 707
Score:	234.50	Matches: 144
Percent Similarity:	32.93%	Conservative: 74
Best Local Similarity:	21.75%	Mismatches: 220
Query Match:	7.54%	Indels: 224
DB:	1	Gaps: 29

US-09-270-437D-5 (1-1708) x FUB2_HUMAN (1-707)

```
Qy   37 AGCCATCATGAAGCTGATGAGTCCACCAGTTGAGAACCATGCCTGAAGGTCTCCTACAT  96
                ||||| :::           |||               |||
Db    2 SerAspTyrSerThrGlyGlyProProGlyProProProProAlaGly---------  18

Qy   97 CCCCATGAGCAGATAGCACAGGGACCTGAGAATGGGCCCGAGGGGCTTTGGCTCTCG  156
                                                     ::: |||
Db   19 ----------------GlyGlyGlyGlyAlaGlyGlyAla  26

Qy  157 GGGTCAGCCCCGCCA---------------------GGGTCCACCTGT  183
            ||||||                              |||  |||
Db   27 GlyGlyGlyProProProGlyAlaGlyAspArgGlyGlyGlyProCys  46

Qy  184 GGCAGCGGGGGCCCCAGCCAAGCAGCAAGTGGA-------CATCCCCCTTCGGCGTCCT  237
            ||||                    ::: |||              ||||||
Db   47 Gly---GlyGlyProGlyGlyGlyGlySerAlaGlyGlyProSerGlnProProGlyGlyGly  65

Qy  238 GGTGCCC--------------------------------ACCCAGTATGTGGG  258
            ||                                    ::: |||
Db   66 Gly-ProGlyIleArgLysAspAlaPheAlaAspAlaValGlnArgAlaArgGlnIleAl  85

Qy  259 TGCCATTATTGGCAAGGAGCGGGGCCCCACCATCCCGCCAACATCACAAAA----------  304
            ||| :::                   ||| |||||
Db   85 aAlaLysIleGlyGlyAspAlaAlaThrThrGlyAspAsnSerThrProAspPheGlyPh  105

Qy  305 ------CAGACCCAGTCCAAGATAGACGTGCATAGGAAGGAGAACGGAGGTGCAGTGA  357
                  ||| ||:::
Db  105 eGlyGlyGlnLysGlyGlnLeuGluAspGluArgGlyAspGluSerLysLysLeuuAlaSe  125

Qy  358 AAAAGCCATCAGTGTGACTCGACTCCACCCCTGAGGCGTGCTCCTCCGCTTGTAAGATGATCTT  417
            :::            |||   :::  |||  |||
Db  125 rGlnGlyAspSerIleSerSerGln------------------------------Le  134

Qy  418 GGAGATTATGCATAAAGAGCTAAGGACCACCAAAAACGGCTGAC---GAGGTTCCCCTGAA  474
            | :::|||                   ::: |||
Db  134 uGlyProIleHisProProProArgThrSerMetThrGluGlyTyrArgValPro-----  152

Qy  475 GATCCTGGCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAGGACGGAACCTGAA  534
            :::      ||||| :::|||||
Db  153 ------AspGlyMetValGlyLeuIleLeuGlyLysGluGluGlyGlyIleIleAs  168

Qy  535 GAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCC----------TCGTTGCA  582
            |||||                :::              ||::::
Db  168 nLysIleGlnIleAspSerGlyCysLysLysValGlnIleSerProAspSerGlyLeu--  187

Qy  583 AGACCTTACCCTTTACAACCCTGAGAGAGACCATCACTGAGAGGGGGCCATCGAGAATTG  642
            ||||||:::::::::::         ||||||
Db  188 -------ProGluArgSerValSerLeuThrGlyAlaProGluSerVa  201

Qy  643 TTGCAGGGCCGAGCAGCAGAAAATAATGAAGAAAGTTCGGGAGGGCCTATGAGAATGATGTGGC  702
            ||:::                              :::
Db  201 ---------AspAspIleVa  212

Qy  703 TGCCATGAGCTTCACCTGATCCTGGCTGCTGCTGTAGGTCTTTTCCCAGC  762
            ||:::                         |||||
Db  212 lSerArgGlyArgGlyGlyProGly---------  221

Qy  763 TTCATCCAGCGCGAGTCCCGCCGCTCCCAGCAGCGTTACTGGGCTGCTCCCTATAGCTC  822
                                         :::
Db  222 ------GlnPheHisAs  225

Qy  823 CTTTATGGAGGCTCCCGAGCAGCAGAGATGGTGCAG--GTGTTTATCCCCGCCAGGCAGT  879
            |||||                              :::
Db  225 pAsnAlaAsnGlyGlyGlnAsnGlyThrValGlnGlyThrIleProAlaGlyLysAl  245

Qy  880 GGGCGCCATCATCGGCAAGAACGGGCAGCACATCAAACAGCTCTCCCGGTTGCCAGCGC  939
            |||||||||||                   :::
Db  245 aGlyLeuValIleGlyGlyGlyGluThrIleLysSerGlnLeuGlnLeuArgAlaGlyVa  265

Qy  940 CTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTT---CGTATGGTTATCAT  996
            |||                 :::        |||||
Db  265 llysMetIleLeuIleGlnLysAspGlySerGlnAsnThrAsnValAspIleSerProLeuArgIl  285
```

```
DE   (Fragment).
GN   NOVA1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BDIX; TISSUE=Cerebellum;
RA   Knudsen A., Monstad S.E., Vedeler C.A.;
RT   "Nova-1, the paraneoplastic Ri antigen, is associated with breast
RT   cancer.";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   FUNCTION.
RX   MEDLINE=22120219; PubMed=12124753;
RA   Kumar D.V., Nighorn A., St John P.A.;
RT   "Role of Nova-1 in regulating alpha2N, a novel glycine receptor
RT   splice variant, in developing spinal cord neurons.";
RL   J. Neurobiol. 52:156-165(2002).
CC   -!- FUNCTION: Functions to regulate alternative splicing in neurons by
CC       binding pre-mRNA in a sequence-specific manner to activate exon
CC       inclusion. It binds specifically to the sequence UCAUY (By
CC       similarity). Acts to regulate a novel glycine receptor alpha-2
CC       chain splice variant (alpha-2N) in developing spinal cord.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Contains 3 KH domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AY262017; AAP20872.1; -.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 3.
DR   SMART; SM00322; KH; 3.
DR   PROSITE; PS50084; KH_TYPE_1; 3.
KW   Antigen; Nuclear protein; RNA-binding; Repeat.
FT   DOMAIN        26     42       BIPARTITE NUCLEAR LOCALIZATION SIGNAL
FT                                 (POTENTIAL).
FT   DOMAIN        48    115       KH 1.
FT   DOMAIN       146    212       KH 2.
FT   DOMAIN       248    384       ALA-RICH.
FT   DOMAIN       396    463       KH 3.
FT   NON_TER      474    474
SQ   SEQUENCE   474 AA;  48443 MW;  70D0926B5BECD131 CRC64;
```

Alignment Scores:

```
Pred. No.:                  1.53e-07              Length:          474
Score:                      231.00                Matches:         106
Percent Similarity:         36.42%                Conservative:    71
Best Local Similarity:      21.81%                Mismatches:      175
Query Match:                7.43%                 Indels:          134
DB:                         1                     Gaps:            15
```

US-09-270-437D-5 (1-1708) x NOA1_RAT (1-474)

```
Qy    125  GAGAATGGGGCGGGCCGAAGGGGGCTTT--------------GGCTCTCGGGGTCAG  163
           :: |||||                            ::: ::|||| |
Db      8  GlnAsnGlyThrHisThrGlyValProIleAspLeuAspProSerArgLysArg---   27

Qy    164  CCCCGGCAGGGCTCACCTGTGCAGCGGGCCCCAGCCAAGCAGCAGCAAGTGGAC---  220
           |||  ::: |||
Db     28  ProLeuGluAlaProProGluAlaGlySerThrLysArgThrAsnThrGlyGluAspGly   47

Qy    221  --ATCCCCTTCGGCTCCGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAG  277
           ||:::||||||:::|| :::                |||:::||||||||||||||
Db     48  GlnTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlySerIleIleGlyLysGly   67
```

RESULT 7

```
NOA1_RAT
ID   NOA1_RAT    STANDARD;    PRT;   474 AA.
AC   Q80WR4;
DT   10-OCT-2003  (Rel. 42, Created)
DT   10-OCT-2003  (Rel. 42, Last sequence update)
DT   10-OCT-2003  (Rel. 42, Last annotation update)
DE   RNA-binding protein Nova-1 (Neuro-oncological ventral antigen 1)
```

The following is the detailed query/database sequence alignment (Qy/Db):

```
Qy    997  CACTCGGACCGCCCAGAGGCCCAATTCAAGGCTCAGGGAAGAATCTATGGCAAAATTCAAGGA  1056
           |||| |||     |||        :::|
Db    285  eIleGlyAspProTyrLysValGlnGlnAlaCysGluMetValMetAspIleLeuArgAs     305

Qy   1057  G-------GTGAAGCTGGAGAC  1098
           :::          ::: |||
Db    305  nValThrLysAlaGlyPheGlyAspArgAsnGluTyrGlyIleGlyGlyIl           325

Qy   1099  CCACATACGTGCCAGCATCAGCAGCTGGCCGGGTCATTGGCAAAGGTGAAAAACGGT  1158
           ::: |||||| ||||||:::  ||||||||:::  :::
Db    325  eAspValProArgHisSerValGlyValValIleGlyArgSerGlyGluMetIl     345

Qy   1159  GAACGAGTTGCAGAATTTGACGGCAGCTGAGGCTGTAGTACCAAGAGACCAGACCCCTGA  1218
           |||  |||
Db    345  eLysLysIleGlnAsn------AspAlaGlyValArgIleGlnPheLysGlnAspAspGl     363

Qy   1219  TGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCTATCGCCAGTCAGATGGCTCA  1278
           :::  ||||||:::::||
Db    363  yThrGlyProGlyLysIleAlaHisIleMetGlyProProAspArgCysGluHisAlaAl     383

Qy   1279  ACGGAAGATCCGAGACATCTCGGCCCAGGTTAAGCAGCAGCATCAGAAGGGACAGAGTAA  1338
           |||  |||  |||:::||
Db    383  aArgIleIleAsnAspLeuLeu------------------------------------     390

Qy   1339  CCAGGCCCAGGCACCGAGGAAGTGACCAGCCCCTCCCTGTCCCTTNGAGTCCAGGACAAC  1398
           |||:::          ||:::
Db    391  ------GlnSerLeuArgSerGlyProProGlyProGly---ProGlyGlyProGlyIlePr     407

Qy   1399  AACGGGCAGAAATGAGAGTGTGCTCTCCCCGGCAGGCCTGAGATGAGTGGGAATCCGG  1458
           |||  |||:::
Db    407  oGlyGlyArgGlyArg------------------GlyArgGlyGlnGlyAsnTrp------     419

Qy   1459  GACACNTGGGCGGGGC------TGTAGATCAGGTTTGCCCACTGATTGAGAAAGATGTT  1512
           ||||||         :::||||
Db    420  ------GlyProGlyGlyGluMetThrPheSerIleProThr-------------Hi     432

Qy   1513  CCAGTGAGGAACCCCTGATCNTNTCAG----------------------------------  1537
           |||  :::
Db    432  sLysCysGlyLeuValIleGlyArgGlyGlyGluAsnValLysAlaIleAsnGlnGlnTh     452

Qy   1538  ----------------CCCCAAACACCCACC--------------  1552
                           |||||||
Db    452  rGlyAlaPheValGluIleSerArgGlnLeuProThrGlyThrSerLysLe     472

Qy   1553  CAATTGGCCCAAACACTGT-------------------  1570
           |||:::
Db    472  uPheIleIleArgGlySerProGlnIleIleAspHisCysArgGlnLeuIleGluGluLuy     492

Qy   1571  --NTGCCCCTCGGGGTGTCAGAAATTNTAGC  1599
           ||:::||||
Db    492  sIleGluGlyProLeuCysProValGlyProGlyProGlyProAlaGlyPr     512

Qy   1600  GCAAGGCACTTTAAACGTGATTGTTTAAAGAAGCT--------------------  1636
           ||| :::     ||  :::
Db    512  oMetGlyProPheAsnProGlyProPheAsnGlnGlyProProGlyAlaProHisAl     532

Qy   1637  ------------CTCCAGGCCCCACCAAGAGGGTGG-----ATCACACTCAGTG  1674
                      ||
Db    532  aGlyGlyProProHisGlnTyrProGlnGlyTyrProGlyAsnThrTyrProGlnTr     552

Qy   1675  GGAA  1678
           |:::
Db    552  pGln  553
```

```
Qy   1088  AAGTTGGAGACCCACATACGTGTGCCAGCATCAGCAGTCGCGGGTCATTGGCAAAGGT  1147
                |||      :::      ||||||       :::||||
Db    400  Glu---------------IleAlaValProGluAsnLeuValGlyAlaIleLeuGlyLysGly  415

Qy   1148  GGAAAAACGTGAACGAGTTGCAGAATTTGACGGCAGCTGAGTGGTAGTACCAAGA---  1204
                |||||||||       |||       :::       |||
Db    416  GlyLysThrLeuValGluTyrGlnLeuLeuThrGlyAlaArgIleGlnIleSerLysLys  435

Qy   1205  ---GACCAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAATCATCGGACATTTCAT  1261
                :::      |||      :::      :::
Db    436  GlyGluProGlyThrArgAsnArgLys---ValThrIleThrGlyThrProAla  454

Qy   1262  GCCAGTCAGATGGCTCAA  1279
                |||||||||
Db    455  AlaThrGlnAlaAlaGln  460


RESULT 8
FUB2_RAT
ID   FUB2_RAT      STANDARD;      PRT;     721 AA.
AC   Q99PF5;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Far upstream element binding protein 2 (FUSE binding protein 2) (KH
DE   type splicing regulatory protein) (KSRP) (MAP2 RNA trans-acting
DE   protein 1) (MARTA1).
GN   FUB2 OR KHSRP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 73-87; 89-108 AND 475-486, SUBCELLULAR
RP   LOCATION, AND FUNCTION.
RC   TISSUE=Brain;
RX   MEDLINE=22246918; PubMed=12358751;
RA   Rehbein M., Wege K., Buck F., Schweizer M., Richter D., Kindler S.;
RT   "Molecular characterization of MARTA1, a protein interacting with the
RT   dendritic targeting element of MAP2 mRNAs.";
RL   J. Neurochem. 82:1039-1046(2002).
CC   -!- FUNCTION: Part of a ternary complex that binds to the downstream
CC       control sequence (DCS) of the pre-mRNA. Mediates exon inclusion in
CC       transcripts that are subject to tissue-specific alternative
CC       splicing. May interact with single-stranded DNA from the far-
CC       upstream element (FUSE). May activate gene expression (By
CC       similarity). Binds to the dendritic targeting element and may play
CC       a role in mRNA trafficking.
CC   -!- SUBUNIT: Part of a ternary complex containing FUBP2, PTBP1, PTBP2
CC       and HNRPH1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear. A small proportion is also found in
CC       the cytoplasm of neuronal cell bodies and dendrites.
CC   -!- SIMILARITY: Contains 4 KH domains.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF308818; AAG59811.1; -.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 4.
DR   SMART; SM00322; KH; 4.
DR   PROSITE; PS50084; KH_TYPE_1; 4.
KW   Transport; mRNA transport; mRNA processing; mRNA splicing;
KW   Transcription regulation; Trans-acting factor; Nuclear protein;
KW   DNA-binding; RNA-binding; Repeat.
FT   DOMAIN      145    209       KH 1.
FT   DOMAIN      234    300       KH 2.
FT   DOMAIN      323    387       KH 3.
```

```
Qy   278  GGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGAGCGTGCATAGGAAG  337
               |||||||       :::      |||      :::
Db    68  GlyGlnThrIleLeuValGlnLeuLysGluThrGlnAlaThrIleLysLeuSerLysSer  87

Qy   338  GAGAAC------GCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGC  391
               :::::       |||       :::::      |||       |||
Db    88  LysAspPheTyrProGlyThrThrGluArgValCysLeuIleGlnIleGlyThrIleGluAla  107

Qy   392  TGCTCCTCCGTTGTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAA  451
               |||::      ||||       :::      :::
Db   108  LeuAsnAlaValHisGlyPheIleAlaGluLysIlePheGlnAlaMetProGlnAsnValAla  127

Qy   452  ACGGCTGACGAGGTTCCCCTG----------  472
              :: :||
Db   128  LysThrGluProValSerIleLeuGlnProGlnThrThrValAsnProAspArgIleLys  147

Qy   473  ------AAGATCCTGCCCCATAATAACTTTGTAGGGCGTCTCATTGGCAAGGAAGGACGG  526
                  |||||:::      |||       :::|||||
Db   148  GlnValLysIleIleIleGluValProAsnSerThrAlaGlyLeuIleIleGlyLysGlyGlyAla  167

Qy   527  AACCCTGAAGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCCTCGTTGCAAGAC  586
               ::||||       |||      :::::
Db   168  ThrValLysAlaIleMetGluGlnSerGlyAlaTrpValGlnIleLeuSerGlnLysProAsp  187

Qy   587  CTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGC  646
               :::      ||||||       |||
Db   188  ---GlyIleAsnLeuGlnLeuArgValValThrValSerGlyValProGluProGlnGlnAsnArg  206

Qy   647  AGGGCCGAGCAGGAGAAATAATGAAGAAAGTTCGGGAG----------  682
              ::||      ||||||::      :::::
Db   207  LysAlaValGluLeuIleIleGluGlnLysIleIleGlnLeuGluAspGlnSerGlySerCysLeu  226

Qy   683  ------GCCATGAGAAT----------  694
               ::|||
Db   227  AsnIleSerTyrAlaAsnValThrGlyProValAlaAsnSerAsnProThrGlyPhePro  246

Qy   695  ------------GATGTGGCTGCCATGAGCTTCCACCTGATCCCTGGC  730
                       |||||      :::::
Db   247  TyrAlaAsnThrAlaGluValLeuProThrAlaAlaAlaAlaAlaGlyLeuLeuGlyHis  266

Qy   731  CTGAACCTGGCTGCTCGTAGGTCTTTTCCCA----------  760
               ||||||||       :::
Db   267  AlaAsnLeuAlaGlyValAlaAlaPheProAlaValLeuSerGlyPheThrGlyAsnAsp  286

Qy   760  ------------------------------  760
Db   287  LeuValAlaIleThrSerAlaLeuAsnThrLeuAlaSerTyrGlyThrTyrAsnLeuAsnThr  306

Qy   761  ------------------GCTTCATCCAGCCAGTCCCGCCTCCCCAGCAGC  796
                           ::||||
Db   307  LeuGlyLeuGlyLeuSerGlnAlaAlaAlaAlaAlaThrGlyAlaAlaLeuAlaAlaAlaAlaSer  326

Qy   797  GTTACTGGGGCTGCTCCCCATAGCTCCCTTATGCAGGCTCCCGAGCAGGAGATGGTGCAG  856
               ||||||       |||||
Db   327  AlaAsnProAlaAlaAlaAlaAlaAsnLeuLeuAlaThrTyrAlaSerGlu---------  343

Qy   857  GTGTTTATCCCGCCCAGGCAGCTGGGCGCCATCATCGGCAAGAGGGGCAGCACATCAAA  916
               ||| |||:::
Db   344  ---------AlaSerAlaSerGlyThrAlaGlyThrAla---------------  355

Qy   917  CAGCTCTCCGGTTTGCCAGGGTTTACACCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCC  976
               |||||||
Db   356  ------GlyThrPheAlaLeuGlySerLeuAlaAlaAla---------------  366

Qy   977  AAAGTTCGTATGGTTATCATCACTGACCCGCCCAATTCAAGGCTCAGGGAGA  1036
               |||
Db   367  ---------ThrAlaAlaThrAsnGlyThrTyrPheGlyAlaSerPro  379

Qy  1037  ATCTATGGCAAACTC----------AAGGAGGAGAACTTCTTTGGTCCCAAGAGGAGGAAGTG  1087
              :::      :::::::
Db   380  LeuAlaAlaSerAlaIleLeuGlyThrGluLysSerThrAspGlySerLysLysAspValVal  399
```

```
FT   DOMAIN      425    492                    KH 4.
FT   DOMAIN      572    685                    4 X 12 AA IMPERFECT REPEATS.
FT   REPEAT      572    583                    1.
FT   REPEAT      618    629                    2.
FT   REPEAT      644    655                    3.
FT   REPEAT      674    685                    4.
FT   DOMAIN       69     68                    GLY/PRO-RICH.
FT   DOMAIN      499    498                    GLY-RICH.
FT   DOMAIN             613                    ALA/GLY/PRO-RICH.
SQ   SEQUENCE    721 AA;   74226 MW;  482C7A765C60E84A CRC64;
```

Alignment Scores:

```
Pred. No.:                 1              Length:          721
Score:                 2.45e-07           Matches:         122
Percent Similarity:     228.50            Conservative:     67
Best Local Similarity:  31.24%            Mismatches:      184
Query Match:            20.17%            Indels:          232
DB:                      7.35%            Gaps:             22
```

US-09-270-437D-5 (1-1708) x FUB2_RAT (1-721)

```
Qy    128  AATGGGCGCCGAGGGGGGCTTTGGCTCTCGGGGTCAGCCCCGCCAG------------------  172
                 ||||  |||||  |||| |||||
Db     98  AsnAsnAsnThrProAspPheGlyPheGlyPheGlyGlyGlnLysArgGlnLeuGluAspGlyAspGly  117
```

```
Qy    173  --------------------------------------------GGCTCACCTGTGGCAGCG  190
                                                        ::: ::|:|||
Db    118  GlnProAspSerLysLeuAlaSerGlnGlyAspSerIleGlySerGlnLeuGlyPro  137
```

```
Qy    191  GGGGCCCCAGGCCAAGCAGCAGCAAGTGACATCCCCTTCGGCTCCTGGTGCCACCCAG  250
                 ||         :::                    ||||||
Db    138  IleHisProProArgThrSerMetThrGluThrGluTyrArg------ValProAspGly  155
```

```
Qy    251  TATGTGGGTGCCATTATTGGCAAGGAGGAGGGGGCCACCATCCCAACATCACAAAACAGACC  310
                 |||||||                ||||:|||:|:|   |||||         ::: :::
Db    156  MetValGlyLeuIleLeuGlyLysGlyValGlyLysGluGlnIleLeuAsnLysLysIleGlnAsnAspSer  175
```

```
Qy    311  CAGTCCCAAGATAGACGTGCATAGGAAGGAAGACAACGCAGGTGCAGTGAAAAAGCCATCAGT  370
                 |||||:: :::    :::::|   |||||||||  ||:||||||
Db    176  GlyCysLysValGlnIle---SerProAspSerGlyLeuProGlnArgSerValSer  194
```

```
Qy    371  GTGCACTCCACCCTGAGGGCGTGCTCCTCGCTTGTAAGATGATCTTGGAGATTATG---  427
                 :::             ||||    ||:: :||:::||:||:
Db    195  LeuThrGlyAlaProGluSerValGlnLysAlaLysMetMetLeuAspAspIleValSer  214
```

```
Qy    428  --------------------------------CATAAAGAGGCTAAGACACCAAAACGGCT  457
                                                ||          :::
Db    215  ArgGlyArgGlyGlyProGlyGlnPheHisSerAspAsnAlaAsnGlyValGlnAsnGly  234
```

```
Qy    458  GACGAGGTGCCTCCCGAGATCCTGCCCCATAATAACTTGTAGGGCGTCTCATTGGCAAG  517
                 ||:|||||||||| |||:|||||||||     ::::|||||| :::|||||||||
Db    235  ThrValGlnGluIleLeuMetIleProAlaGlyHisLys------AlaGlyLeuValIleLeuGlyLysLys  252
```

```
Qy    518  GAAGGACGGAACCTGAAGAAGGTAGAGCAAGAATACCGAGACAAAAATCACCATC------  571
                 ||:||||||||||:||||||| |||  |||     |||:||
Db    253  GlyGlyGluThrIleLeuLysGlnLeuGlnLeuGlyArgAlaGlyValLysMetIleLeuGln  272
```

```
Qy    572  TCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGAGGACCATCACTGTGAAGGGGGCC  631
                 ||||:::   :: ||||||||||       ||        ||||||||
Db    273  AspGlySerGlnAsnThrAsnValAspLysProLeuArgIleIleGlyAspProTyrLys  292
```

```
Qy    632  ATCGAGAATTGTTGCAGGCCCAGCCCAGACAGCAGAAATAATGAAGAAAGTTCGGGAGAG  691
                 :::::      |||       :::||||||||::::::
Db    293  ValGlnGlnAlaCysGluMetValMetAspIleLeuArg---------------GluLeu  306
```

```
Qy    692  AATGATGTGGCTGCCATGAGCTCTCACCTGATCCCTGGCCTGGCTGCTGTAGGT-----ATCACACTT  751
                 ||                             CTCCAGGCCCCACCAAGAGGGTGG
Db    307  ArgAspGlnGlnGlyPheGlyPheAspArgArgAsnGluTyrGly--------------  319
```

```
Qy    752  CTTTTCCCAGCTTCATCCAGCCAGTCCCGCCGCTTCCCGGCTCGTTACTGGGGGCTGCT  811
                 |||  |||  ||    ||||||
Db    320  --------------------------------------------SerArgValGlyGlyGly---  325
```

```
Qy    812  CCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGCAGGAGATGGTCAGGTGTTTATCCCGCC  871
                 :::  ||||||
Db    326  ------------------------------------------IleAspValProValProArg  332
```

```
Qy    872  CAGGCAGTGGGCGCCATACTGGCAAGAGAGGGGCAGCACATCAAACAGCTCTCCGGTTT  931
                 :::|||||||||||||||   :::||||||||||||||||||||
Db    333  HisSerValGlyValValIleGlyTyrArgSerGlyMetIleLeuLysLysIleGlnAsnAsp  352
```

```
Qy    932  GCCAGCGCGCCTCCATC-------------------------------------AAGAATTGCA  955
                 |||                                              ||||||
Db    353  AlaGlyValAlaArgIleGlnPheLysGlnAspGlyThrGlyProGluLysIleAlaHis  372
```

```
Qy    956  -------CCACCCGAAACACCTGACTCACTCCAAAGTTCGTATGGTT-----------  991
                 ||||||||||:::      :::
Db    373  IleMetGlyProAspArgCysGluHisAlaAlaArgIleIleAsnAspLeuLeuGln  392
```

```
Qy    992  --ATCATCACTGACCGCCAGAGGCCCAATT--------------------------  1020
                 :::    ::|||||||||||||| |||||
Db    393  SerLeuArgSerGlyValProPro---GlyProProGlyAlaProGlyMetProProGlyGlyA  412
```

```
Qy   1021  ----CAAGGCTCAGGGAAGAAATCATGGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCC  1075
                 ::: |||  |||  |||||:::  |||||||
Db    412  rgGlyArgGlyArgGlyGlyGlnGlyLysAsnTrpGlyLysGluMetThrPhe-----  430
```

```
Qy   1076  AAGGAGGAAGTGAAGCTGGAGACCCACATACGTGCCAGCCATCAGCAGCTGGCCGGGTC  1135
                 ::: ::||||
Db    431  ------------------------------------------Ser-IleProThrHisLysCysGlyLeuVal  440
```

```
Qy   1136  ATTGGCAAAGGTGGAAAAACGGTGAACGGATTGACGGCAGTGAGGTGGTA  1195
                 |||||||:::|| ||||||||:|||||||||||||
Db    441  IleGlyArgGlyGlyGlyGlyLysAsnValLysAlaIleAsnGlnGlnThrGlyAlaPheValGlu  460
```

```
Qy   1196  GTACCAAGAGACCAGACCCCT------GATGAGAACGACCAGGTCATCGTG--------  1240
                 |||||||||||||                |||
Db    461  IleSerArgGlnLeuProProAsnGlyAspProAsnGlyLeuProAsnPheLysAspGlyArgGly  480
```

```
Qy   1241  ----AAAATCATCGGACATTTCTATGCCAGTCAGATGCTCAACGGAAGATCCGAGAC---  1294
                 ::|        ||||||
Db    481  SerProGlnGlnIleGluAspHis------AlaLysGlnLeuIleGluGlnLysLysIleGlu---  497
```

```
Qy   1295  ATCCTGGCCCCAGGTAAGCAGCAGCATCGAAGAGGACAGAGTAACCAGGCCCAGGCACGG  1354
                 |||
Db    497  ------------------------------------------------------------  497
```

```
Qy   1355  AGGAAGTGACCAGCCCCTCCCTGTCCCTTGAGTCCAGGACCACCAAAAACGGCTCAGAAATCGA  1414
                 |||||:::  :::|||||||
Db    498  --------------------------------GlyProLeuCysPro  502
```

```
Qy   1415  GAGTGTGCTCTCCCCGCAGCCTGAGAATGAGTGGGAATCGGACACNTGGGCGGGC  1474
                 |||||||||      ::::||||
Db    503  ------------------------------------------------ValGlyProGly  506
```

```
Qy   1475  TGTAGATCAGGTTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCTGATCTNT  1534
                 |||||||:::  :::
Db    506  ------------------------------------------------------------  506
```

```
Qy   1535  CAGCCCCAAACACCCACCCAATTGGCCCACACTGTNTGCCCTCGGGGTGTCAGAAATT  1594
                 ||||||||:::  :::
Db    507  ------------------------------ProGlyGlyProGlyProAla  513
```

```
Qy   1595  NTAGGCCAAGGCACTTTAAACGTCGATTGTTTAAAGAAGCT-----------------  1636
                 ::||||||||       :::
Db    514  GlyProMetGlyProPheHisProGlyProPheAsnGlnGlyProProGlyAlaProPro  533
```

```
Qy   1637  ------------------------------CTCCAGGCCCCACCAAGAGGGTGG------ATCACACTT  1669
                                                |||  |||:::
Db    534  HisAlaAlaGlyGlyProProProHisGlnTyrProProGlnTyrProProGlnThrTyrPro  553
```

```
Qy   1670  CAGTGGGAA  1678
                 |||:::
Db    554  GluTrpGln  556
```

RESULT 9

```
PCB1_HUMAN
ID   PCB1_HUMAN      STANDARD;      PRT;   356 AA.
AC   Q15365; Q11157; Q14975;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly(rC)-binding protein 1 (Alpha-CP1) (hnRNP-E1) (Nucleic acid
DE   binding protein SUB2.3).
GN   PCBP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95331278; PubMed=7607214;
RA   Kiledjian M., Wang X., Liebhaber S.A.;
RT   "Identification of two KH domain proteins in the alpha-globin mRNP
RT   stability complex.";
RL   EMBO J. 14:4357-4364(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96016208; PubMed=7556077;
RA   Leffers H., Dejgaard K., Celis J.E.;
RT   "Characterisation of two major cellular poly(rC)-binding human
RT   proteins, each containing three K-homologous (KH) domains.";
RL   Eur. J. Biochem. 230:447-453(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymphocytes;
RX   MEDLINE=94203810; PubMed=8152927;
RA   Aasheim H.-C., Loukianova T., Deggerdal A., Smeland E.B.;
RT   "Tissue specific expression and cDNA structure of a human transcript
RT   encoding a nucleic acid binding (oligo(dC)) protein related to the
RT   pre-mRNA binding protein K.";
RL   Nucleic Acids Res. 22:959-964(1994).
CC   -!- FUNCTION: Single-stranded nucleic acid binding protein that binds
CC       preferentially to oligo dC.
CC   -!- SUBCELLULAR LOCATION: Loosely bound in the nucleus. May shuttle
CC       between the nucleus and the cytoplasm.
CC   -!- TISSUE SPECIFICITY: Abundantly expressed in skeletal muscle,
CC       thymus and peripheral blood leucocytes while a lower expression is
CC       observed in prostate, spleen, testis, ovary, small intestine,
CC       heart, liver, adrenal and thyroid glands.
CC   -!- PTM: PHOSPHORYLATED. THE NON-PHOSPHORYLATED FORM(S) EXHIBITED THE
CC       STRONGEST POLY(RC)-BINDING ACTIVITY.
CC   -!- SIMILARITY: Contains 3 KH domains.
CC   -!- CAUTION: There is probably a frameshift error in the nucleotide
CC       sequence of Ref.3.
```

```
DR   EMBL; X78137; CAA55016.1; -.
DR   EMBL; U24223; AAA91317.1; -.
DR   EMBL; 229505; CAA82631.1; ALT_FRAME.
DR   HSSP; Q07244; 1KHM.
DR   Genew; HGNC:8647; PCBP1.
DR   GK; Q15365; -.
DR   MIM; 601209; -.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0005634; C:nucleus; NAS.
DR   GO; GO:0003723; F:RNA binding; IDA.
DR   GO; GO:0003697; F:single-stranded DNA binding; IDA.
DR   GO; GO:0016071; P:mRNA metabolism; NAS.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 3.
DR   SMART; SM00322; KH; 3.
DR   PROSITE; PS50084; KH_TYPE_1; 3.
KW   Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;
KW   Phosphorylation; Repeat.
FT   DOMAIN       13     75       KH 1.
FT   DOMAIN       97    162       KH 2.
FT   DOMAIN      279    343       KH 3.
FT   CONFLICT    205    205       V -> A (IN REF. 2 AND 3).
SQ   SEQUENCE   356 AA;  37526 MW;  DC854775776C5104 CRC64;
```

```
Alignment Scores:
Pred. No. :                        5.92e-07        Length:            356
Score:                             221.50          Matches:            76
Percent Similarity:                41.78%          Conservative:       74
Best Local Similarity:             21.17%          Mismatches:        138
Query Match:                        7.12%          Indels:             71
DB:                                 1               Gaps:                9
```

```
US-09-270-437D-5 (1-1708) x PCB1_HUMAN (1-356)

Qy     188 GCGGGGGCCCCAGCCAAGCAAGCAGCAAGTGGACATCCCCCCTTCGGCTCCTGGTGCCACC 247
               ||||||                                     ::: :||||||:::
Db       3 AlaGlyValThrGluSerGlyLeuAsnValThrLeuThrIleArgLeuLeuMetHisGly 22

Qy     248 CAGTATGTGGGTGCCATTATTGGCAAGAGAGGGGCCACCATCCGCAACATCACAAAACAG 307
               :::  |||||:||:::||||||||||:||:||||                :||:||::|||
Db      23 LysGluValGlySerIleIleGlyLysGluSerValLysGlySerValIleArgGluGlu 42

Qy     308 ACCCAGTCCAAGATAGACGTGCATAGGAGAGGAAGGAAGGAACGCAGGTGCAGCTGAAAAAGCCATC 367
               ::: :::|||||:|::::|                                ||||:|||
Db      43 SerGlyIleAlaArgIleAsnIle------SerGluGlyAsnCysSerProGluArgIleIle 59

Qy     368 AGTGTGCACTCCACCCCTGAGGGCTGCTCTCCGGTTGTAAGATGATCTTGGAGATTATG 427
               :::::                          ||||||||::||::::::::::||
Db      60 ThrLeuThrGlyProThrAsnAlaIlePheLysAlaPheAlaMetIleIleAspLysLysLeu 79

Qy     428 CATAAAGAGGCTAAGGACCACCAAAACG---------GCTGACGAGGTTCCC 469
               ::::::                ||| |||
Db      80 GluGluAspIleAsnSerSerMetThrAsnSerThrAlaSerArgProProValThr 99

Qy     470 CTGAAGATCCTGGCCCATAATAACTTTGTAGGGCGTCATTGGCAAGGAAGGACGGAAC 529
               ||:::::::::::::::::              |||    |||||||||||||||
Db     100 LeuArgLeuValValProAlaThrGlnCysGlySerIleuIleGlyLysGlyGlyCysLys 119

Qy     530 CTGAAGAAGGTAGGAGCAAGATACCGACAGAAAATCACCATCTCCTGTTGCAAGACCTT 589
               ::|||||||| :::  ||| :: |||||||                ||||:|||:||
Db     120 IleLysGluIleArgGluSerThrGlyAlaGlnValGlnValAlaGly-----AspMet 137

Qy     590 ACCCTTTACAACCCTGAGAGACCATCACTGTGAGAGGGGCCATCGAGAATTGTTGCAGG 649
               ::        ||||||||||            |||           |||
Db     138 LeuProAsnSerThrGluThrGlnAlaIleThrIleAlaGlyValProGlnSerValThrGlu 157

Qy     650 GCCGAGCAGGAAATA--------------ATGAAGAAGTTCGGGAGGCCTATGAGAATGAT 697
               ::::::||||                :::|||   ::: :::::: :::
Db     158 CysValLysGlnIleCysLeuValMetLeuGluThrLeuSerGlnSerProGlnArg 177

Qy     698 GTGGCTGCCATGAGCTCACCTGATCCCTGGCCTGAACCTGGCT------------ 742
               ||        ||| ||||
Db     178 ValMetThrIleProTyrGlnProMetProAlaSerSerProValIleCysAlaGlyGly 197

Qy     743 ------------------GCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCCGCCG 784
                                 ||||:||| :::::|||||||
Db     198 GlnAspArgCysSerAspAlaValGly---TyrProHisAlaThrHisAspLeuGluGlyGly 216

Qy     785 CCTCCC------------------------------ 790
               ::::::
Db     217 ProProLeuAspAlaTyrSerIleGlnGlyGlnHisThrIleSerProLeuAspLeuAla 236

Qy     790 ------------------------------------ 790

Db     237 LysLeuAsnGlnValAlaArgGlnIleSerHisPheAlaMetMetHisGlyThrGly 256

Qy     791 ---AGCAGCGTTACTGGGCGTGCTCCC-----------------TATAGCTCCTTTATGCAGGGCT 835
```

```
Db   257  PheAlaGlyIleAspSerSerProGluValLysGlyTyrTrpAlaSerLeuAspAla  276
          ::: :::        |||    :: :::       |||  :::       |||
Qy   836  CCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCAGGCAGTGGGCGCCATCATCGGC  895
          ||||  ||||| ||      |||||     :: ||| ||||||      ||||||||
Db   277  SerThrGlnThrThrHisGluLeuThrIleProAsnAsnLeuIleGlyCysIleIleGly  296

Qy   896  AAGAAGGGGCAGCACCATCAAACAGCTCTCCGGTTGCCAGCGCCTCCATCAAGATTGCA  955
          ::::::||||   :::|||  :::::: ::  :::    |||  |||||||||||||||
Db   297  ArgGlnGlyAlaAsnIleAsnGluIleAsnGlnMetSerGlyAlaGlnIleLeuLysIleAla  316

Qy   956  CCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCACTGGACCGGCCAGAGGCC  1015
          ||                                       ::: |||
Db   317  AsnPro---ValGluGlySerSerGlyArgGlnValThrIleThrGlySerAlaAlaSer  335

Qy   1016 CAATTCAAGGCTCAGGGAAGAAATCATGGCAAACTCAAGGAGGAGAACTCTTTGGT  1072
          ||||| |||            ||         :::||||   |||
Db   336  IleSerLeuIleGlnTyrLeuIleAsnAlaArgLeuSerGluSerGluLysGlyMetGly  354

RESULT 10
NOA1_MOUSE
ID   NOA1_MOUSE      STANDARD;      PRT;     493 AA.
AC   Q9JKN6; Q8C8B9;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   RNA-binding protein Nova-1 (Neuro-oncological ventral antigen 1)
DE   (Ventral neuron-specific protein 1) (Fragment).
GN   NOVA1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CBA/J; TISSUE=Brain;
RX   MEDLINE=20458856; PubMed=11003693;
RA   Ward-Bailey P.F., Wood B., Johnson K.R., Bronson R.T., Donahue L.R.,
RA   Davisson M.T.;
RT   "Neuromuscular ataxia: a new spontaneous mutation in the mouse.";
RL   Mamm. Genome 11:820-823 (2000).
RN   [2]
RP   SEQUENCE OF 250-493 FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Wells C.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573 (2002).
RN   [3]
RP   FUNCTION.
RX   MEDLINE=20181810; PubMed=10719891;
RA   Jensen K.B., Dredge B.K., Stefani G., Zhong R., Buckanovich R.J.,
RA   Okano H.J., Yang Y.Y.-L., Darnell R.B.;
RT   "Nova-1 regulates neuron-specific alternative splicing and is
RT   essential for neuronal viability.";
RL   Neuron 25:359-371 (2000).
RN   [4]
RP   FUNCTION.
RX   MEDLINE=22692957; PubMed=12808107;
RA   Dredge B.K., Darnell R.B.;
RT   "Nova regulates GABA(A) receptor gamma2 alternative splicing via a
RT   distal downstream UCAU-rich intronic splicing enhancer.";
RL   Mol. Cell. Biol. 23:4687-4700(2003).
CC   -!- FUNCTION: Functions to regulate alternative splicing in neurons by
CC       binding pre-mRNA in a sequence-specific manner to activate exon
CC       inclusion. It binds specifically to the sequence UCAUY. Most
CC       likely acts to activate the inclusion of exon E3A in the glycine
CC       receptor alpha-2 chain and of exon E9 in gamma-aminobutyric-acid
CC       receptor gamma-2 subunit via a distal downstream UCAU-rich
CC       intronic splicing enhancer.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DISEASE: Defects in NOVA1 leads to neuronal death in spinal and
CC       brainstem neurons.
CC   -!- SIMILARITY: Contains 3 KH domains.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; AF22828; AAF35907.1; -.
DR   EMBL; AK047565; BAC33088.1; -.
DR   MGD; MGI:104297; Nova1.
DR   GO; GO:0005634; C:nucleus; IC.
DR   GO; GO:0003729; F:mRNA binding; IDA.
DR   GO; GO:0006371; P:mRNA splicing; IDA.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 3.
DR   SMART; SM00322; KH; 3.
DR   PROSITE; PS50084; KH_TYPE_1; 3.
KW   Antigen; Nuclear protein; RNA-binding; Repeat.
FT   NON_TER       1      1
FT   DOMAIN       13     29                       (POTENTIAL).
FT   DOMAIN       35    102       KH 1.
FT   DOMAIN      157    223       KH 2.
FT   DOMAIN      259    395       ALA-RICH.
FT   DOMAIN      407    474       KH 3.
SQ   SEQUENCE   493 AA;  50302 MW;  7D09E8A55B0A7817 CRC64;

Alignment Scores:
Pred. No.:                    6.38e-07          Length:             493
Score:                        221.50            Matches:            107
Percent Similarity:           35.63%            Conservative:        69
Best Local Similarity:        21.66%            Mismatches:         167
Query Match:                  7.12%             Indels:             151
DB:                           1                 Gaps:                16

US-09-270-437D-5 (1-1708) x NOA1_MOUSE (1-493)

Qy   152  TCTCGGGGTCAGCCCCGCCAGGGCTCACCTGTGCCAGCGGGCGGGGGCCCAGCCAAGCAGCAGCAG  211
          ||| ||||||    :::||||| :::||||||  ||   :::
Db    11  SerArgLysArgProLeuGluAlaProGluAlaGlySerThrLysThrArgThrAsnThr  30

Qy   212  CAAGTGGAC-----ATCCCCCTTCGGTCTCCGGCTCCGGTGCCCACCAGTATGTGGGTGCCATT  265
          |||              |||:::::||:::||||:::||||   |||::::|||
```

```
Db    31 GlyGluAspGlyGlnTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlySerIle 50

Qy   266 ATTGGCAAGGAGGGGGCCACCATCCAAAACAGACCCAGTCCAAGATAGAC 325
         ||||||||||      :::       ||:|| ||||:
Db    51 IleGlyLysGlyGlyGlnThrIleValGlnLeuGlnLysGlyGlnGluThrIleAlaThrIleLys 70

Qy   326 GTGCATAGGAAGGAGGAAC------GCAGGTGCAGTCGAGCCATCAGTGTGCCACTCC 379
         :::   :::::         |||||||| ||     :::        ||||||
Db    71 LeuSerLysSerLysAspPheTyrProGlyThrThrGlnValCysLeuIleGlnIleGly 90

Qy   380 ACCCCTGAGGCGTGCTCCTCCGCT---------------TGTAAGATGATCTTGGAGATTATG 427
         :::      |||       |||:::
Db    91 ThrIleGluAlaLeuAsnAlaValHisGlyPheIleAlaGluLysLysIleArgGluMetPro 110

Qy   428 CATAAAGAGGCTAAG----- 442
         |||||
Db   111 GlnAsnValAlaLysThrGluProValSerIleLeuGlnProGlnThrThrValAsnPro 130

Qy   443 ----------------------------GACACCAAAACGGCTGACGAGGTT 466
                                        ||||||:::::
Db   131 AspArgIleLysGlnThrLeuProSerSerProThrThrLysSerSerProSerAsp 150

Qy   467 CCCCTG----------------AAGATCCTGGCCCATAAACTTTGTAGGG 502
         ||||||             ||    |||||
Db   151 ProMetThrThrSerArgAlaAsnGlnValLysIleIleValProAsnSerThrAlaGly 170

Qy   503 CGTCTCATTGGCAAGGAAGACGGAACCTGAGAAGGTAGAGCAAGATACCGAGACAAAA 562
         |||||||||||        :::||  || ||| :::
Db   171 LeuIleIleGlyLyLysGlyGlyAlaThrValLysAlaThrValLysAlaIleMetGluGlnSerGlyAlaTrp 190

Qy   563 ATCACCATCTCCTCGTTGCAAGACCTTACCCTTTACAACCCTGAGGAGACCACTGTG 622
         :::  :::|||          |||   :::|||||||
Db   191 ValGlnLeuSerGlnLysProAsp--GlyIleAsnLeuGlnLysGlnArgValValThrVal 209

Qy   623 AAGGGGGCCATCGAGAATTGTTGCAGAGGGCCGAGCAGGAAATAATGAAGAAGTTCGGGAG 682
         ||||||        :::|||      ||||||||||:::|||||||||
Db   210 SerGlyGluProGluGlnAsnArgLysAsnAlaValGluLeuIleGlnLysLysIleGlnGlu 229

Qy   683 --------------------GCCATGAGAAT--- 694
                              :::|||  |||
Db   230 AspProGlnSerGlySerCysLeuAsnIleSerTyrAlaAlaAsnValThrGlyProValAla 249

Qy   695 ---------------------------------GATGTGGCTGCC 706
                                            |||||||
Db   250 AsnSerAsnProThrGlySerProTyrAlaAlaAsnThrAlaGluValLeuProThrAlaAla 269

Qy   707 ATGAGGCTCTCACCTGATCCCTGGCCTGGACCTGGCTGTGTAGGTCTTTTCCCA------ 760
         :::|||||  ||||:||:||  |||||       |||
Db   270 AlaAlaAlaGlyLeuLeuGlyHisAlaAsnLeuAlaGlyValAlaLeuAlaPheProAlaVal 289

Qy   760 -------------------------------------------------GCTTCATCCAGC 772
                                                          |||||
Db   290 LeuSerGlyPheThrGlyAsnAspLeuAsnAspLeuThrSerAlaLeuAsnThrLeuAsnAla 309

Qy   761 --------------------------------------------------GCTTCATCCAGC 772
Db   310 SerTyrTyrGlyTyrAsnLeuAsnThrLeuGlyLeuGlyLeuGlySerGlnAlaAlaAlaThrGly 329

Qy   773 GCAGTCCCGCGCCGCCCTCCAGCAGCGTTACTGGGGCTGCCCCTATAGCTCCTTTATGCAG 832
         |||:::  |||          ||||||:::
Db   330 AlaLeuAlaAlaAlaAlaAlaSerAlaAsnProAlaAlaAlaAlaAlaAsnLeuAlaAla 349

Qy   833 GCTCCCGAGCAGCAGGAGATGTGCAGGTGTTTATCCCGCCAGTGGGCGCCATCATC 892
         ||||||        |||
Db   350 ThrTyrAlaSerGlu-----------------------AlaSerAlaSerGlySerThrAla 362

Qy   893 GGCAAGAAGGGCAGCACATCAAACAGCTCTCCCGGTTTGCCAGCGCCTCCATCAAGATT 952
         |||||                    :::||||||
Db   363 GlyGlyThrAla-------------------GlyThrPheAlaLeuGlySerLeuAlaAla 376

Qy   953 GCACCACCCGAAACACCTGACTCCTCCAAAGTTCGTATGTTATCATCACTGGACCGCCAGAG 1012
         |||
Db   377 Ala-------------------------------ThrAlaAlaThrAsn 382

Qy  1013 GCCCAATTCAAGGTCAGGAAGAAGAATCTATGGCAAACTC---------AAGGAGGAGAAC 1063
         |||                :::
Db   383 GlyTyrPheGlyAlaGlyAlaAlaSerProLeuAlaAlaSerAlaIleLeuGlyGluGlyLysSer 402

Qy  1064 TTCTTTGGTCCCAAGGAGGAGAGTGAAGGTGAGACCGACCCACATACGTGTCCAGCATCAGCA 1123
         :::       :::           ||    :::
Db   403 ThrAspGlySerLysAspValValGlu---------IleAlaValProGluAsnLeu 418

Qy  1124 GCTGGCGCCGGTCATTGGCAAAGGTGGCAAAAACCGTGAACGAGTTGCAGAATTGACGGCA 1183
         ||        :::::||||||||||:::||||||||||
Db   419 ValGlyAlaIleAlaLeuLeuGlyLysGlyGlyLysThrLeuValGlyThrGlnLeuGluLeuUThrGly 438

Qy  1184 GCTGGAGGTGGTAGTACCAAGA-------GACCAGACCCCTGATGAGAACGACCAGGTCATC 1237
         ||                      :::      :::::
Db   439 AlaArgIleGlnIleLeuSerLysGlyGlyUPheValProGlyThrArgArgAsnArgLys--- 457

Qy  1238 GTGAAAATCATCCGACATTTCATGCCAGTCAGATGGCTCAA 1279
         |||:::||
Db   458 ValThrIleThrGlyThrProAlaAlaThrGlnAlaAlaGln 471

RESULT 11
NOA1_HUMAN
ID   NOA1_HUMAN             STANDARD;      PRT;     510 AA.
AC   P51513;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   RNA-binding protein Nova-1 (Neuro-oncological ventral antigen 1)
DE   (Onconeural ventral antigen-1) (Paraneoplastic Ri antigen) (Ventral
DE   neuron-specific protein 1).
GN   NOVA1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC   TISSUE=Cerebellum, and Hippocampus;
RX   MEDLINE=94000830; PubMed=8398153;
RA   Buckanovich R.J., Posner J.B., Darnell R.B.;
RT   "Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding
RT   protein and is specifically expressed in the developing motor
RT   system.";
RL   Neuron 11:657-672(1993).
RN   [2]
RP   SEQUENCE OF 1-34 FROM N.A.
RC   TISSUE=Fetal brain;
RA   Dmitrenko V.V., Garifulin O.M., Shostak K.A., Smikodub A.I.,
RA   Kavsan V.M.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 423-495.
RX   MEDLINE=99148126; PubMed=10368286;
RA   Lewis H.A., Chen H., Edo C., Buckanovich R.J., Yang Y.Y.-L.,
RA   Musunuru K., Zhong R., Darnell R.B., Burley S.K.;
RT   "Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding
RT   domains.";
RL   Structure 7:191-203(1999).
CC   -!- FUNCTION: May regulate RNA splicing or metabolism in a specific
CC       subset of developing neurons.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=1;
CC         IsoId=P51513-1; Sequence=Displayed;
CC       Name=2; Synonyms=Tumor;
CC         IsoId=P51513-2; Sequence=VSP_002841, VSP_002843, VSP_002844;
CC       Name=3;
CC         IsoId=P51513-3; Sequence=VSP_002842;
CC   -!- TISSUE SPECIFICITY: Brain.
CC   -!- DISEASE: Autoantigen in the paraneoplastic opsoclonus myoclonus
CC       ataxia (POMA), a paraneoplastic neurological syndrome/disorder
```

```
CC       (PNS/D) associated with breast cancer, fallopian cancer, and SCLCa
CC       and characterized primarily by loss of inhibitory control of motor
CC       neurons in the spinal cord and brainstem. Recognized by the IgG
CC       autoantibody ANNA-2 (also called anti-Ri).
CC   -!- SIMILARITY: Contains 3 KH domains.
CC   ---------------------------------------------------------------------
CC       This SWISS-PROT entry is copyright. It is produced through a collaboration
CC       between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC       the European Bioinformatics Institute. There are no restrictions on its
CC       use by non-profit institutions as long as its content is in no way
CC       modified and this statement is not removed. Usage by and for commercial
CC       entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC       or send an email to license@isb-sib.ch).

DR   EMBL; U04840; AAA16022.1; -.
DR   EMBL; Z70771; CAA94810.1; -.
DR   PIR; I38489; I38489.
DR   PDB; 1DT4; 10-MAY-00.
DR   Genew; HGNC:7886; NOVA1.
DR   GK; P51513; -.
DR   MIM; 602157; -.
DR   GO; GO:0003723; F:RNA binding; TAS.
DR   GO; GO:0007626; P:locomotory behavior; TAS.
DR   GO; GO:0008380; P:RNA splicing; TAS.
DR   GO; GO:0007268; P:synaptic transmission; TAS.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 3.
DR   SMART; SM00322; KH; 3.
DR   PROSITE; PS50084; KH_TYPE_1; 3.
KW   Antigen; Nuclear protein; RNA-binding; Repeat; Alternative splicing;
KW   3D-structure.
FT   DOMAIN      27     43       BIPARTITE NUCLEAR LOCALIZATION SIGNAL
FT                               (POTENTIAL).
FT   DOMAIN      49    119       KH 1.
FT   DOMAIN     174    240       KH 2.
FT   DOMAIN     276    412       ALA-RICH.
FT   DOMAIN     424    491       KH 3.
FT   VARSPLIC    88     90       Missing (in isoform 2).
FT                               /FTId=VSP_002841.
FT   VARSPLIC   153    176       Missing (in isoform 3).
FT                               /FTId=VSP_002842.
FT   VARSPLIC   177    184       VKIIVPNS -> KHNISWIS (in isoform 2).
FT                               /FTId=VSP_002843.
FT   VARSPLIC   185    510       Missing (in isoform 2).
FT                               /FTId=VSP_002844.
SQ   SEQUENCE   510 AA;  52055 MW;  7B9BA8B67CAB6984 CRC64;
```

Alignment Scores:

```
Pred. No.:             6.43e-07       Length:             510
Score:                 221.50         Matches:            111
Percent Similarity:    34.89%         Conservative:        68
Best Local Similarity: 21.64%         Mismatches:         173
Query Match:           7.12%          Indels:             161
DB:                    1              Gaps:                17
```

US-09-270-437D-5 (1-1708) x NOA1_HUMAN (1-510)

```
Qy    338 GAGAACGCA-----------------------------GGTGCAGCTGAAAAGCCATCAGTGTGCACTCCACC 382
              |||                                  |||  :::
Db     89 SerLysSerLysAspPheTyrProGlyThrThrGluArgValCysLeuIleGlnGlyThr 108

Qy    383 CCTGAGGGCTGCTCCTCCCGCT--------------------TGTAAGATGATCTTGGAGATTATGCAT 430
              :::                               :::  ||| |||:::
Db    109 ValGluAlaLeuAsnAlaValHisGlyPheIleAlaGluLysIleArgGluMetProGln 128

Qy    431 AAAGAGGCTAAG--------------------------------------------------- 442
              ||||||
Db    129 AsnValAlaLysThrGluProValSerIleLeuGlnProGlnThrThrValAsnProAsp 148

Qy    443 --------------------------------GACACCAAAACGGCTGACGAGGTTCCC 469
                                              |||||||:::
Db    149 ArgIleLysGlnThrLeuProSerProThrThrLysSerProSerProSerAspPro 168

Qy    470 CTG------------------------AAGATCCTGGCCCATATAACTTTGTAGGGGGT 505
              :::                      |||:::||
Db    169 MetThrThrSerArgAlaAsnGlnHisValLysIleIleValProAsnSerThrAlaGlyLeu 188

Qy    506 CTCATTGGCAAGGAAGGAGGACGGAACCTGAAGAAGTGAGCAGAAGTACCGAGACAAAAATC 565
              :::||||||  :::|| |||:::::
Db    189 IleIleGlyLysGlyAlaThrValLysAlaValMetGluGlnSerGlyAlaTrpVal 208

Qy    566 ACCATCTCCGTTGCAAGACCTTACCCTTACAACCCTGAGAGGACCATCACTGTGAAG 625
              ||||  :::||||:::|||
Db    209 GlnLeuSerGlnLysProAsp---GlyIleAsnLeuGlnLysGlnArgValThrValSer 227

Qy    626 GGGGCCATGAGAATTGTTGCAGGCCAGAATGAGAAAGTTCGGGAG-- 682
              ||  :::  :::|||:::|||||:::
Db    228 GlyGluProGluGlnAsnArgLysAlaValGluLeuIleGluGlnLysGlnGluAsp 247

Qy    683 -----------------------GCCTATGAGAAT-- 694
                                     ::|||
Db    248 ProGlnSerGlySerCysLeuAsnIleSerTyrAlaAsnValThrGlyProValAlaAsn 267

Qy    695 ------------------------------GATGTGGCTGCCATG 709
                                            ||||
Db    268 SerAsnProThrGlySerProTyrAlaAsnThrAlaGluValLeuProThrAlaAlaAla 287

Qy    710 AGTCTCACCTGATCCTGGCCTGAACCTGGCTGCTGCTGTAGGTCTTTTCCCA------- 760
              :::::: ||||:::||  |||:::
Db    288 AlaAlaGlyLeuLeuGlyHisAlaAsnLeuAlaLeuAlaGlyValAlaAlaAlaPheProAlaValLeu 307

Qy    760 ---------------------------------GCTCATCCAGCGCA 775
                                               |||
Db    308 SerGlyPheThrGlyAsnAspLeuValAlaAlaIleThrSerAlaLeuAsnThrLeuAlaSer 327

Qy    761 ---------------------------------GCTCATCCAGCGCA 775
                                               |||  :::
Db    328 TyrGlyTyrAsnLeuAsnThrLeuGlyLeuGlyLeuSerGlnAlaAlaIleAlaThrGlyAla 347

Qy    776 GTCCCCGGCCTCCCAGCAGCGTTACTGGCGTGCTCCCATAGCTCCTTTATGCAGGCT 835
              |||||        :::|||
Db    348 LeuAlaAlaLeuAlaAlaSerAlaAsnProAlaAlaAlaIleAlaAsnLeuAlaThr 367

Qy    836 CCCGAGCAGGAGGATGGTGCAGGTGTTTATCCCGCGCCAGGCAGTGGGCGCCATCATCGGC 895
              |||||:::
Db    368 TyrAlaSerGlu---------------------AlaSerAlaSerGlySerThrAlaGly 380

Qy    896 AAGAAGGGGCAGCAATCAAACAGCTCTCCCGGTTGCCAGCGCCTCCATCAAGATTGCA 955
              |||||
Db    381 GlyThrPheAla---------------------GlyThrPheAlaLeuGlySerLeuAlaAlaAlaAla 394

Qy    956 CCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCATGACCGCCAGAGGCC 1015
              |||||||:::
Db    395 ---------------------ThrAlaAlaThrAsnGly 400

Qy   1016 CAATTCAAGGCTCAGGGAAGAAATCATGGCAAACTC--------AAGGAGGAGAACTTC 1066
              ||| :::
Db    401 TyrPheGlyAlaAlaSerProLeuAlaAlaSerAlaIleLeuGlyThrGluLysSerThr 420

Qy    125 GAGAATGGGCGCCGCAGGGGGCTTT------------------------GGCTCTCGGGGTCAG 163
              :::||||||||                                 ||||| :::
Db      9 GlnAsnGlyThrHisThrGlyValProIleAspLeuAspProSerArgLysArg 28

Qy    164 CCCCGCCAGGGCTCACCTGCAGCGGGAGCCCCAGCCAGCAGCCAGCAAGTGAC---220
              ||| :::                                              |||
Db     29 ProLeuGluAlaProProGluAlaGlySerThrLysArgThrAsnThrArgGlyHisAspGly 48

Qy    221 --ATCCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAG 277
              ||:::::|||||||||:::
Db     49 GlnTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlyIleLeuGlyLysGly 68

Qy    278 GGGGCCACCATCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGGAAG 337
              ||||||||:::  ||:::
Db     69 GlyGlnThrIleValGlnGlnLeuGlnLysGlnThrGlyAlaThrIleLysLeuSerLysLeu 88
```

```
Qy    1067 TTTGGTCCCAAGGAGGAAGTGAAGCTGGAGAGACCCACATACGTGCCAGCATCAGCAGCT 1126
                |||  |||::: ||||::          |||  ||||||||||
Db     421 AspGlySerLysAspValValGlu--------IleAlaValProGluAsnLeuVal 436

Qy    1127 GGCCGGGTCATTGGCAAAGGTGGAAAAACGGTGAACGAGTTGCAGAATTTGACGGCAGCT 1186
                ||||||||::::||||||||||||||||||:::||
Db     437 GlyAlaIleLeuGlyLysGlyGlyLysThrValGluTyrGlnGluLeuThrGlyAla 456

Qy    1187 GAGGTGTAGTACCAAGA------GACCAGACCCCTGATGAGAACGACCAGGTCATCGTG 1240
                :::  |||  :::  :::        ||||
Db     457 ArgIleGlnIleSerLysLysGlyGluPheValProGlyThrArgAsnArgLys---Val 475

Qy    1241 AAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAA 1279
                |||||||||||::||  ||||||
Db     476 ThrIleThrGlyThrProAlaAlaThrGlnAlaAlaGln 488
```

RESULT 12
```
PCB1_RABIT
ID   PCB1_RABIT    STANDARD;    PRT;   356 AA.
AC   O19048;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly(rC)-binding protein 1 (Alpha-CP1) (hnRNP-E1).
GN   PCBP1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9207122; PubMed=10101190;
RA   Thiele B.J., Berger M., Huth A., Reimann I., Schwarz K., Thiele H.;
RT   "Tissue-specific translational regulation of alternative rabbit
RT   15-lipoxygenase mRNAs differing in their 3'-untranslated regions.";
RL   Nucleic Acids Res. 27:1828-1836(1999).
CC   -!- FUNCTION: Single-stranded nucleic acid binding protein that binds
CC       preferentially to oligo dC (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: Contains 3 KH domains.
CC   --------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------
DR   EMBL; AJ003023; CAA05814.1; -.
DR   HSSP; Q07244; 1KHM.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 3.
DR   SMART; SM00322; KH; 3.
DR   PROSITE; PS50084; KH_TYPE_1; 3.
KW   Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding; Repeat.
FT   DOMAIN        13     75       KH 1.
FT   DOMAIN        97    162       KH 2.
FT   DOMAIN       279    343       KH 3.
SQ   SEQUENCE   356 AA;  37497 MW;  6D1A261276CA206D CRC64;
```

Alignment Scores:
```
Pred. No.:                          Length:             356
Score:                   218.50     Matches:             75
Percent Similarity:      41.90%     Conservative:        75
Best Local Similarity:   20.95%     Mismatches:         139
Query Match:              7.03%     Indels:              69
DB:                            1    Gaps:                 9
```

```
US-09-270-437D-5 (1-1708) x PCB1_RABIT (1-356)

Qy     188 GCGGGGGCCCCAGCCAAGCAGCAGCAAGTGGACATCCCCCCTTCGGCTCCTGGTGCCCACC 247
```

```
Db       3 AlaGlyValThrGluSerGlyLeuAsnValThrLeuThrIleArgLeuLeuMetHisGly 22
                ||||||

Qy     248 CAGTATGTGGTGGCCATTATTGGCAAGGAGGGGGGCCACCATCCGCAACATCACAAACAG 307
                :::  ||||||||::::||||||||||||||||||||:::||
Db      23 LysGluValGlySerIleIleGlyLysGlyGlySerValIleSerArgIleArgGluGlu 42

Qy     308 ACCCAGTCCAAGATAGACGTGCATAGGAACGAGGAGAACGACCAGGTGCAGTGAAAAGCCATC 367
                :::::::  :::||||||  :::         |||
Db      43 SerGlyAlaArgIleAsnIle--------SerGluGlyAsnCysProGluArgIleIle 59

Qy     368 AGTGTGCACTCCACCCCTGAGGGCTCCTCCCGCTGTAAGATGATCTTGGAGATTATG 427
                :::::            ||||||  ||::||  |||||||:::||
Db      60 ThrLeuThrGlyProThrAsnAlaIlePheLysAlaPheAlaMetIleIleAspLysLeu 79

Qy     428 CATAAAGAGCTAAGGACACCAACAAACG----------GCTGACGAGGTTCCC 469
                :::::            :::  ||||||
Db      80 GluGluAspIleAsnSerSerMetThrAsnSerThrAlaAlaSerArgProProValThr 99

Qy     470 CTGAAGATCCTGGCCCCATAATAACTTTGTAGGGCGTTCCATTGGCAAGGAAGGACGGAAC 529
                :::  ||||||||||  ||||||||||||||||
Db     100 LeuArgLeuValValProAlaThrGlnIleCysLeuValMetGlyLysGlyGlyLysCysLys 119

Qy     530 CTGAAGAAGGTAGGAGCAAGATACCGAGACAAAAAATCACCACCCTCGTTGCAAGACCTT 589
                :::||||::::::  :::  ||||||  :::|||
Db     120 IleLysGluIleArgGluSerThrGlyAlaGlnValGlnValAlaGly------AspMet 137

Qy     590 ACCCTTTACACCCTGAGGAGACCATCACTGTGAAGGGGGCCATCGAGAATTGTTGCAGG 649
                :::  |||||||||:::||||||  |||
Db     138 LeuProAsnSerThrGluArgAlaAlaIleThrIleAlaGlyValProGlnSerValThrGlu 157

Qy     650 GCCGGACGAGGAAATA----------ATGAAGAAAGTTCGGGAGGCCTATGAGAATGAT 697
                :::||||||         |||  :::
Db     158 CysValLysGlnIleCysLeuValMetLeuGlyThrLeuSerGlnSerProGlnGlyArg 177

Qy     698 GTGGCTGCCATGAGCTCTCCACCTGATCCCTGGCCTGAAC----------CTGGCTGCTGTAGGT 751
                |||         |||  :::         :::  :::|||
Db     178 ValMetThrIleProTyrGlnProMetProAlaSerSerProValIleCysAlaGlyGly 197

Qy     752 CTT--------------------------------------TTCCAGCTTCATCCAGCGGCAGTCCCGCGCCT 787
                |||                                       :::
Db     198 GlnAspArgCysSerAspAlaAlaAlaGlyTyrProHisAlaThrHisAspLeuGluGlyProGlyPro 217

Qy     788 CCC----------------------------------------------------TATAGCTCCTTTATCCAGGCTCCC 790
                |||
Db     218 ProLeuAspAlaTyrSerIleGlnInGlnHisThrIleSerProLeuAspLeuAlaLys 237

Qy     790 ------------------------------------------------------ 790
                
Db     238 LeuAsnGlnValAlaAlaArgGlnGlnSerHisPheAlaMetMetHisGlyThrGlyPhe 257

Qy     791 AGCAGCGTTACTGGGGCTGCTCCC------------------------------TATAGCTCCTTTATCCAGGCTCCC 838
                :::  :::  ||||||
Db     258 AlaGlyIleAlaSerSerSerProGluValLysGlyTyrTrpProAlaSerLeuAspAlaSer 277

Qy     839 GAGCAGGAGGATGTGCAGTGTTTATCCCGCCCACCAGCCAGTGGCGGCCATCATCGGCAAG 898
                ||||         :::::            ||||||||||||||
Db     278 ThrGlnThrThrHisGluLeuThrHisGluProAsnAsnLeuIleGlyCysIleIleGlyArg 297

Qy     899 AAGGGGCCAGCCACCATCAAACAGCTCTCCGGTTTGCCAGCCGCTCCCATCAAGATTGCACCA 958
                ||||||||||||||||||:::||  :::
Db     298 GlnGlyAlaAsnIleAsnGluIleArgGlnMetSerGlyAlaGlnIleLysLysIleAlaAsn 317

Qy     959 CCCGAAACACCTGACTCCAAGTTCGTATGGTTATCATCACTGGACCTGGCGGAGGCCCAA 1018
                |||||                 :::::::::||||||
Db     318 Pro---ValGluGlySerGlyArgGlnValThrIleThrGlySerAlaAlaSerIle 336

Qy    1019 TTCAAGGCTCAGGGAAGAAATCATGGCAAACTCAAGGAGGAGAACTTCTTTGGT 1072
                |||              ||||||||||||||:::
Db     337 SerLeuAlaGlnThrTyrLeuLeuIleAsnAlaArgLeuSerGlyLeuSerGluLysGlyMetGly 354
```

RESULT 13
PCB4_HUMAN

ID   PCB4_HUMAN   STANDARD;   PRT;   403 AA.
AC   P57723;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Poly(rC)-binding protein 4 (Alpha-CP4).
GN   PCBP4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20396135; PubMed=10936052;
RA   Makeyev A.V., Liebhaber S.A.;
RT   "Identification of two novel mammalian genes establishes a subfamily
RT   of KH-domain RNA-binding proteins.";
RL   Genomics 67:301-316(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA   Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA   Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA   Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA   Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Hopkins R.F., Jordan H., Farmer A.A., Rubin G.M., Hong L.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Single-stranded nucleic acid binding protein that binds
CC   preferentially to oligo dC (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: Contains 3 KH domains.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF176330; AAG09241.1; -.
DR   EMBL; AK023993; BAB14761.1; -.
DR   EMBL; BC003008; AAH03008.1; -.
DR   EMBL; BC004153; AAH04153.1; -.
DR   Genew; HGNC:8652; PCBP4.
DR   GO; GO:0005634; C:nucleus; ISS.
DR   GO; GO:0003677; F:DNA binding; ISS.

DR   GO; GO:0003723; F:RNA binding; NAS.
DR   GO; GO:0016071; P:mRNA metabolism; ISS.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 3.
DR   SMART; SM00322; KH; 3.
DR   PROSITE; PS50084; KH_TYPE_1; 3.
KW   Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;
KW   Repeat.
FT   DOMAIN   17   67   KH 1.
FT   DOMAIN   101   154   KH 2.
FT   DOMAIN   241   293   KH 3.
SQ   SEQUENCE   403 AA;   41481 MW;   3D99F762A9471265 CRC64;

Alignment Scores:
Pred. No.:   403
Score: 1.11e-06   217.50
Percent Similarity: 42.63%   80
Conservative: 53
Best Local Similarity: 25.64%   114
Query Match: 6.99%   65
Indels: 11
DB: Gaps: 1

US-09-270-437D-5 (1-1708) x PCB4_HUMAN (1-403)

```
Qy    215  GTGGACATCCCCCCTTCGGCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAG  274
               ::|       |||||:|||       |||||||||:::||:|
Db     16  IleThrLeuThrLeuArgMetLeuMetHisGlyTyrAsnGlyLeuValGlySerIleIleGlyIleGlyLys  35

Qy    275  GAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGACGTGCATAGG  334
               :::||    |||:|||     |||       |||||||     :||:||
Db     36  LysGlyGluThrValLysArgIleArgGluGlnSerAlaArgIle--------------  51

Qy    335  AAGGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCTGAGGGCTGC  394
               ||||       |||||
Db     52  ------------------------------ThrIleSerGluGlySerCysProGluArgIle  62

Qy    395  TCCTCCGCTGTAAGATGATCTTGGAGATTATGCATAAAGAGCTAAGGACACCAAAACG  454
               :::       |||       ::: ||
Db     63  ThrThrIleThrGlySerThrAlaAlaValPheHisAlaAlaValSerMetIleAlaAlaPheLys  82

Qy    455  GCTGACGAG-------------------------------------------------GTT  466
               ||:||       |||||       |||:
Db     83  LeuAspGluCysAlaAlaProLeuAlaAlaProLysAsnGlyAsnValSerArgProVal  102

Qy    467  CCCCTGAAGATCCTGGCCCCCATAATAACTTTGTAGGCGTCCATTGGCAGGAGGAAGGACGG  526
               |||       :::       |||||||||
Db    103  ThrLeuArgLeuValIleProAlaSerGlnCysGlySerLeuIleGlyLysValAlaGlyThr  122

Qy    527  AACCTGAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCATTCCTCCGTTGCAAGAC  586
               :::||:|||       ::: |||       |||       :::::
Db    123  LysIleLysGluGlyIleArgGluThrThrGlyAlaGlnValGlnValAlaGly------Asp  140

Qy    587  CTTACCCTTACAACCCTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTGTTGC  646
               :::|||       |||||       ::: 
Db    141  LeuLeuProAsnSerThrGluArgAlaValAlaValThrValSerGlyValProAspAlaIleIle  160

Qy    647  AGGGGCCGAGCAGGAAATAATGAAGAAAGTTCGGAGGCCATGAGCATATGATGTGGCTGCC  706
               ||:::
Db    161  LeuCysValArgGlnIleCysAlaValIleLeuGluSerProLysGly---AlaThr  179

Qy    707  ATGAGCTCCACCTGATCCTGGCCTGAACCTGGCTGCTGTA------------------------  748
               :::       |||       |||||:::||
Db    180  IleProTyrHis------ProSerLeuSerLeuGlyThrValLeuLeuSerAlaAsnGln  197

Qy    749  -----------------------------GGTCTTTTCCCAGCT------------------  763
                                            |:||||
Db    198  GlyPheSerValGlnGlyGlnTyrGlyAlaValThrProAlaGluValThrLysLeuGln  217

Qy    764  ------TCATCCAGCGAGTCCG---CCGCCTCCCAGCAGCGTTACTGGGCGCTGCTCCC  814
                     |||||||       |||||       ||:::
Db    218  GlnLeuLeuSerSerHisAlaValProPheAlaThrProSerValProGlyLeuAspPro  237

Qy    815  TATAGCTCCTTTATGCAGGCTCCCGACAGGAGATGGTGCAGGAGATGGTGTTTATCCCCGCCCAG  874
```

```
Qy    215  GTGGACATCCCCTCGGCTCCTCGTGCCACCCAGTATGTGGTGCCATTATTGGCAAG  274
                :::  |||||||||  ::: ||| |||||||||||  |:|||||||
Db     16  IleThrLeuThrLeuArgMetLeuMetHisGlyLysLysGlyGluValGlySerIleIleGlyGlyLys   35

Qy    275  GAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAGATAGAGCGTGCATAGG  334
                ::: |||||||||  ::: ::::::|||
Db     36  LysGlyGluThrValLysArgIleArgGluGlnSerSerAlaArgIle----------  51

Qy    335  AAGGAGAACGCAGGTGCAGCTGAAAAAGCCATCAGTGTTGCACTCCACCCTGAGGGGTGC  394
                                      ---ThrIleSerGluGlyGlySerCysProGluArgIle  62
Db     52  ------------------------

Qy    395  TCCTCGCCTTGTAGATGATCTTGGAGATTATCGATAAAGAGGCTAAGGACACCACCAAAACG  454
                :::::::
Db     63  ThrThrIleThrGlySerThrAlaAlaValPheHisAlaValSerMetIleAlaPheLys   82

Qy    455  GCTGACGAG----------------------------------------------GTT  466
                |||||
Db     83  LeuAspGluCysAlaAlaProAlaAsnGlyGlySerValSerArgProProVal  102

Qy    467  CCCCTGAAGATCCTGGCCCATAATAACTTGTAGGGCGTCATTGGCAAGAAGGACGG  526
                |||||  :::    :::::::
Db    103  ThrLeuArgLeuValIleProAlaSerGlnCysGlySerLeuIleGlyLysAlaGlyThr  122

Qy    527  AACCTGAAGAAGGTAGAGCAAGAGCAAGAACAAAATCACCATCTCCGTTGCAAGAC  586
                :::|||||:|||  :: :||   :::::::  |||
Db    123  LysIleLysGluIleLysGluThrThrGlyAlaGlnValGlnValAlaAlaGly------Asp  140

Qy    587  CTTACCCTTTACAACCCTGAGACGGACCATCACTGTGAGGGGGCCATCGAGAATTGTGTGC  646
                :::
Db    141  LeuLeuProAsnSerThrGluValArgAlaValThrValSerGlyValProAspAlaIleIle  160

Qy    647  AGGGCCGAGCAGGAAATAATGAAGAAAGTTCGGGAGGCTCATGAGAATGATGTGGCTGCC  706
                :::::|||
Db    161  LeuCysValArgGlnIleCysAlaValIleLeuGluSerProLysGly--AlaThr  179

Qy    707  ATGAGCTCTCACCTGATCCTGGCCTGAACCTGGCTGCTGTA----------------------  748
                :::  |||
Db    180  IleProTyrHis------ProSerLeuSerLeuGlyThrValLeuLeuSerAlaAsnGln  197

Qy    749  ----------------------------------------GGTCTTTTCCCAGCT-------  763
Db    198  GlyPheSerValGlnGlyGlnThrGlyValAlaValThrProAlaGluValThrLysLeuGln  217

Qy    764  ----------TCATCCAGCGCAGTCCG---CCGCTCCCAGCAGCCTTACTGGGGCTGTCCCC  814
                          ||||                |||   |||
Db    218  GlnLeuSerGlyHisAlaValProPheAlaSerProSerValProGlyMetAspPro  237

Qy    815  TATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCCAG  874
                                                     :::|||
Db    238  SerThr----------GlnThrSerSerGlnGluPheLeu------ValProAsnAsp  251

Qy    875  GCAGTGGGCGCCATCATCGGCAAGAAGGGGCAGCACCATCAAACAGCTCTCCCGGTTTGCC  934
                :::|||||||||||  :::  |||
Db    252  LeuIleGlyCysValIleGlyArgGlnGlyGlySerLysIleSerGluIleArgGlnMetSer  271

Qy    935  AGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATC  994
                                      :::
Db    272  GlyAlaHisIleLysIleGly--AsnGlnAlaGluGlyAlaGlyHisValThr  290

Qy    995  ATCACTGACCCGCCAGAGGCCCAATTCAAGGCTCAG  1030
                |||||||||  :::
Db    291  IleThrGlySerProValSerIleAlaLeuAlaGln  302
```

RESULT 14
PCB4_MOUSE
ID   PCB4_MOUSE   STANDARD;   PRT;   403 AA.
AC   P57724;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Poly(rC)-binding protein 4 (Alpha-CP4).
GN   PCBP4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20396135; PubMed=10936052;
RA   Makeyev A.V., Liebhaber S.A.;
RT   "Identification of two novel mammalian genes establishes a subfamily
RT   of KH-domain RNA-binding proteins.";
RL   Genomics 67:301-316(2000).
CC   -!- FUNCTION: Single-stranded nucleic acid binding protein that binds
CC   preferentially to oligo dC (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- TISSUE SPECIFICITY: Ubiquitous.
CC   -!- SIMILARITY: Contains 3 KH domains.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AF176328; AAG09239.1; -.
DR   MGD; MGI:1890471; Pcbp4.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 3.
DR   SMART; SM00322; KH; 3.
DR   PROSITE; PS50084; KH_TYPE_1; 3.
KW   Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;
KW   Repeat.
FT   DOMAIN   17   67   KH 1.
FT   DOMAIN   101   154   KH 2.
FT   DOMAIN   241   293   KH 3.
SQ   SEQUENCE   403 AA; 41398 MW; D4DBFF490CEFE90A CRC64;

Alignment Scores:
Pred. No.:
Score: 1.73e-06
Percent Similarity: 214.50
Best Local Similarity: 42.31%
Query Match: 25.32%
DB: 6.90%

	Length:	403
	Matches:	79
	Conservative:	53
	Mismatches:	115
	Indels:	65
	Gaps:	11

US-09-270-437D-5 (1-1708) x PCB4_MOUSE (1-403)

```
Db    238  GlyThr--------GlnThrSerSerGlnGluPheLeu------ValProAsnAsp  251
                :::  |||                                 :::|||
Qy    875  GCAGTGGGCGCCATCATCGGCAAGAAGGGGCAGCACCATCAAACAGCTCTCCCGGTTTGCC  934
                :::|||||||||||  :::  |||
Db    252  LeuIleGlyCysValIleGlyArgGlnGlyGlySerLysIleSerGluIleArgGlnMetSer  271
                          |||
Qy    935  AGCGCCTCCATCAAGATTGCACCACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATC  994
                                        :::  |||
Db    272  GlyAlaHisIleLysIleGly--AsnGlnAlaGluGlyAlaGlyHisValThr  290
                |||||  :::
Qy    995  ATCACTGACCCGCCAGAGGCCCAATTCAAGGCTCAG  1030
                |||||||||  :::
Db    291  IleThrGlySerProValSerIleAlaLeuAlaGln  302
```

RESULT 15
Y475_ARATH
ID   Y475_ARATH   STANDARD;   PRT;   606 AA.
AC   P58223; O49507; Q8H0Y7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Putative nucleic acid binding protein At4g18375.
GN   AT4G18375 OR F28J12.2.

```
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.;
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purmelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sehkon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA   Nelson J., Spieth J., Ryan B., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaugnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   STRAIN=cv. Columbia;
RX   MEDLINE=22954850; PubMed=14593172;
RA   Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA   Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA   Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA   Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA   Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama Y., Ansari Y.,
RA   Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA   Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA   Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA   Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA   Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA   Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA   Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT   "Empirical analysis of transcriptional activity in the Arabidopsis
RT   genome.";
```

```
RL   Science 302:842-846(2003).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P58223-1; Sequence=Displayed;
CC         Note=No experimental confirmation available;
CC       Name=2;
CC         IsoId=P58223-2; Sequence=VSP_008899, VSP_008900;
CC         Note=May be due to a competing acceptor site. No experimental
CC         confirmation available;
CC   -!- SIMILARITY: Contains 5 KH domains.
CC   -!- CAUTION: Ref.1 sequences differ from that shown due to erroneous
CC       gene model prediction. At4g18370 and At4g18375 were originally
CC       fused into a single gene.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AL021710; CAA16717.1; ALT_SEQ.
DR   EMBL; AL161548; CAB78839.1; ALT_SEQ.
DR   EMBL; AY133701; AAM91635.1; -.
DR   EMBL; BT001108; AAN64172.1; -.
DR   InterPro; IPR004087; KH_dom.
DR   InterPro; IPR004088; KH_type_1.
DR   Pfam; PF00013; KH; 5.
DR   SMART; SM00322; KH; 5.
DR   PROSITE; PS50084; KH_TYPE_1; 5.
KW   Hypothetical protein; Nuclear protein; RNA-binding; Repeat;
KW   Alternative splicing.
FT   DOMAIN        35     99       KH 1.
FT   DOMAIN       138    210       KH 2.
FT   DOMAIN       311    380       KH 3.
FT   DOMAIN       394    455       KH 4.
FT   DOMAIN       535    599       KH 5.
FT   VARSPLIC     532    532       L -> F (in isoform 2).
FT                                 /FTId=VSP_008899.
FT   VARSPLIC     533    606       Missing (in isoform 2).
FT                                 /FTId=VSP_008900.
SQ   SEQUENCE   606 AA;  65760 MW;  61F135BB8647C0C CRC64;
```

```
Alignment Scores:
Pred. No.:                         4.32e-06                    Length:          606
Score:                             209.00                      Matches:         105
Percent Similarity:                37.62%                      Conservative:     91
Best Local Similarity:             20.15%                      Mismatches:      201
Query Match:                        6.72%                      Indels:          124
DB:                                 1                          Gaps:             22

US-09-270-437D-5 (1-1708) x Y475_ARATH (1-606)

Qy      230  CGGCTCCTGGTGCCCACCCAGTATGTGGGTGCCATTATTGGCAAGGAGGGGGCCACCATC  289
             ||:::||| |||                    |||||| ::|||||||||||||||
Db       39  ArgIleLeuCysProIleAspValValGlyValIleGlyLysSerGlyLysLysValIle   58

Qy      290  CGCAACATCACAAAACAGACCCAGTCCAAGCAGCTGCATAGAGAGGAGGAACGGCAGGT   349
             |||                |||||||:||||||||| |||        ::: |||
Db       59  AsnAlaIleArgHisAsnThrLysAlaLysIleLysValPheAspGlnLeuHis--Gly   77

Qy      350  GCAGCTGAAAAAGCCATCAGTGTGCAC-------------------------------  376
             ::::::::||:||::::|||
Db       78  CysSerGlnArgValIleThrIleTyrCysSerValLysGluLysGluGluIleGly   97

Qy      377  -------TCCACCCCTGAGGGCTGCTCCTCCGCTTGTAAGATGATCTTGGAG------  421
                    ||||||||                       :: ||||:
Db       98  PheThrLysSerGluAsnGluProLeuCysAlaGlnAspAlaLeuLeuLysValTyr  117

Qy      422  -----------ATTATGCATAAAGAGGCTAAGGACACCACCAAAACG----------GCTGAC  460
```

Db 458 ArgLeuArgGluAspValLeuGlyAspLysAspSerValAlaThrArgLys------ 474
          |||
Qy 1367 GCCCCTCCCTGTCCCTTNGAG-----TCCAGGACAACACGGGCAGAAATCGAGAGTGT 1420
     |||||| ::: |||
Db 475 ---ProProAlaArgThrAspAsnCysSerPheLeuSerGlySerSerAsnAlaGlyTyr 493

Qy 1421 GCTCTTCCCC----------------------------GGC 1432
     |||||||
Db 494 ThrLeuProSerPheMetSerSerMetAlaSerThrSerGlyPheHisGlyTyrGlySer 513

Qy 1433 AGGGCTGAGAATGAGTGGGAATCCGGGACACNTGGGCCGGGCTGTAGATCAGGTTGCCC 1492
     ::: |||||| |||||
Db 514 PheProAlaGlyAspAsnValLeuGlySerThrGlyLeuGlySerThrGlyProTyrSerTyrGlyArgLeuPro 533

Qy 1493 ACT 1495
     :::
Db 534 Ser 534

Search completed: July 13, 2004, 12:09:27
Job time : 42.5 secs

Db 118 AspAlaAlaSerAspGluGluAsnAsnThrLysThrAsnValAspArgAspAsp 137
     :::   :::||||||||||
Qy 461 GAGGTTCCCTGAAGATCCTGGCCCAATAATAACTTTGTAGGCGTTCATTGGCAAGGAA 520

Db 138 AsnLysGluCysArgLeuLeuValProPheSerGlnSerSerSerLeuIleGlyLysAla 157
     :::||:::
Qy 521 GGACGGAACCTGAAGAGAGTAGACAAGATACGCAGACAAAAATCACCATCTCCTCGTTG 580

Db 158 GlyGluAsnIleLysArgIleLeuArgArgThrArgAlaSerValLysValValSer--- 176
     ||| |||:::||:::||||
Qy 581 CAAGACCTTACCCTTTACAACCT----------------------GAGAGGACCATC 616

Db 177 LysAspValSer-----AspProSerHisValCysAlaMetGluTyrAspAsnValVal 194
     :::||::: ::: |||
Qy 617 ACTGTGAAGGGG--------GCCATCGAGAATTGTTGCAGGGCCGAGCAGGAAATAATG 667

Db 195 ValIleSerGlyGluProGluSerValLysGlnAlaLeuPheAlaValSerAlaIleMet 214
     |||:::
Qy 668 AAGAAAGTTCGG------------GAGGCCTATGAGAATGATGTGGCTGCC 706

Db 215 TyrLysIleAsnProArgGluAsnIleProLeuAspSerThrSerGlnAspValProAla 234
     :::::::||||||
Qy 707 ATGAGC-------------- 712

Db 235 AlaSerValIleValProSerAspLeuSerAsnSerValTyrProGlnThrGlyPheTyr 254

Qy 713 --------TCTCACCTG---------------ATCCCTGGC---CTGAACCTGGCT 742
     ||:::   :::||
Db 255 SerAsnGlnAspHisIleLeuIleGlnGlnGlyAlaGlyValIleProSerTyrPheAsnAlaLeu 274

Qy 743 GCTGTAGGTCTTTCCCA-----GCTTCATCCAGCGCAGTCCCCGCCCTCCCAGCAGC 796
     :::||| :::||| :::
Db 275 SerValSerAspPheGlnGlyTyrArgAlaGluThrAlaAlaAsnProValProValPheAla 294

Qy 797 GTTACTGGGCGCTGCTCCCTATAGCTCCTTTATGCAGGCTCCCGAGCAGGAGATGGTG--- 853
     ::: :::||||||:::
Db 295 SerSerLeuProValThrHisGlyPheGlyGlySerArgSerGluGluValPhe 314

Qy 854 CAGGTGTTTATCCCGCCAGGCAGTGGGCGCCATCATCGGCAAGAAGGGCAGCACATC 913
     :::|| ::: ||||||:::
Db 315 LysValLeuCysProLeuCysAsnIleMetArgValIleGlyValIleGlyLysGlySerThrIle 334

Qy 914 AAACAGCTCTTCCCGGTTGCCAGCGGCCTCCATCAAGATTGCACCACCCGAAACACCTGAC 973
     :::|||:::|| :::
Db 335 LysArgIleArgGluAlaSerGlyCysIle-------------GluValAsnAsp 349

Qy 974 TCCAAAGTTCGT-----------ATGGTTATCATCACTGGACCGCCGCCAGAGGCC 1015
     ||:::   :::||||||:::
Db 350 SerArgThrLysCysGlyAspAspAspCysValIleIleIleValThrArgAlaThrGluSerPro 369

Qy 1016 CAATTCAAGGCTCAGGGAGGAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCC 1075
     :::
Db 370 AspAspMetLysSerMetAlaValGluAlaValLeuLeuLeuLeuGlnGluTyrIleAsnAsp 389

Qy 1076 AAGGAGGAAGTGAAGCTGGAGACCCACATACGTGCCAGCATCAGCAGCTGCCGGGTC 1135
     :::||| ::: |||
Db 390 GluAspAlaGluAsnValLysMetGlnLeuLeuLeuValSerSerLysIleGlyCysVal 409

Qy 1136 ATTGGCAAAGGTGGAAAAACGGTGAAGAGTTGCAGAATTGACGGCAGCTGAGGTGGTA 1195
     ||||||::: :::||:::
Db 410 IleGlyLysSerGlySerValIleIleAsnGluIleLeuArgThrAsnAlaAsnIleCys 429

Qy 1196 GTACCAAGAGACCAAGACCCCTGATGAGAACGACCAGGTCATCGTGAAAAATCATCGGACAT 1255
     ||| ::::::: :::|||
Db 430 IleSerLysGlyLysLysLysLysAspAsp-----------LeuValGluValSerGlyGlu 444

Qy 1256 TTCTCATGCCAGTCAGATGGCTCAACGGAGATCCGAGAGATCCGAGACATCCTGGCCCAGTT----- 1309
     :::||| |||::: |||
Db 445 Val-----------------SerSerValArgAspAlaLeuIleGlnIleValLeu 457

Qy 1310 --AAGCAGCAGCATCAGAAGGGACAGAGTAACCAGGCCCAGGTAACCGAGGAAGTGACCA 1366
     ::: |||:::